Space, self, and the theater of consciousness

Arnold Trehub

Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA

Abstract

Over a decade ago, I introduced a large-scale theory of the cognitive brain which explained for the first time how the human brain is able to create internal models of its intimate world and invent models of a wider universe. An essential part of the theoretical model is an organization of neuronal mechanisms which I have named the Retinoid Model [Trehub, A. (1977). Neuronal models for cognitive processes: Networks for learning, perception and imagination. *Journal of Theoretical Biology*, *65*, 141–169; Trehub, A. (1991). *The Cognitive Brain*: MIT Press]. This hypothesized brain system has structural and dynamic properties enabling it to register and appropriately integrate disparate foveal stimuli into a perspectival, egocentric representation of an extended 3D world scene including a neuronally tokened locus of the self which, in this theory, is the neuronal origin of retinoid space. As an integral part of the larger neuro-cognitive model, the retinoid system is able to perform many other useful perceptual and higher cognitive functions. In this paper, I draw on the hypothesized properties of this system to argue that neuronal activity within the retinoid structure constitutes the phenomenal content of consciousness and the unique sense of self that each of us experiences.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Consciousness; Retinoid system; Egocentric space; Self; Attention; Illusion; Perspectival experience; Cognitive brain

1. Introduction

The foundational premise of this paper is summed up as follows:

Our cognitive brain is especially endowed with neuronal mechanisms that can model within their biological structures all conceivable worlds, as well as the world we directly perceive or know to exist. External expressions of an unbounded diversity of brain-created models constitute the arts and sciences and all the artifacts and enterprises of human society (Trehub, 1991).

If this claim is true, two important implications follow. First, there must be a biophysical structure in the brain with properties which constitute our internal aspects of self in the world and warrant the lexical designation "self" in philosophical and scientific argument as well as in common discourse. Second, there is a more inclusive system of mechanisms which, in their encounters with real and imagined worlds, determine all that we can possibly say about consciousness.

E-mail address: trehub@psych.umass.edu

By what biologically plausible means can the human brain create the diversity of veridical and hypothetical models/representations claimed above? I argue that (1) there is a system of mechanisms in the brain whose minimal properties can be specified and which can be demonstrated to have the competence to create such models, (2) that these neuronal mechanisms provide an internal structural analog of 3D space within which all phenomenal events are represented, and (3) that the system utilizes a compact and dynamic neuronal center which represents the self within the brain's analog of 3D space. This paper will support Strawson's contention that the self is a thing (Strawson, 1997), and elaborate on what the self-thing might be and what it might do.

In the following sections, I hope to convey an understanding of how the structure and dynamics of a particular model of the cognitive brain (Trehub, 1991) constrain and conform to some conspicuous aspects of our perceptual experience. Section 2 briefly discusses some key aspects of human adaptation.

Section 3 is a summary of my theoretical model of the human brain's embodiment of space (the neuronal structure and dynamics of what I call the retinoid system), while supporting evidence for the operative details of the retinoid model is presented in Section 4. Sections 5 and 6 relate the neuronal model to the concepts of self and space, and present a new explanation of the centuries-old puzzle of the moon illusion as evidence in support of the retinoid model of egocentric space. Section 7 deals with the interaction of analogical and symbolic representation, while Section 8 argues that the mechanisms and systems of the extended neuronal brain model provide a credible biological grounding for the metaphorical notion of a theater of consciousness within the cognitive brain.

2. The need for an egocentric representation of 3D space

We cannot imagine our own behavior without an awareness at the same time of a 3D space around us within which our behavior occurs. One of the distinguishing characteristics of humans is the ability to think of future situations and to plan ahead. We not only make mental representations of the existing layout of salient objects in our immediate spatial environment, but we are also able to "picture" the structure of hypothetical environments which may contain different objects in different locations. The self is the indispensable object that must be represented in our sensible space when we adapt our behavior to current circumstances, when we contemplate our behavioral choices, and when we plan for future contingencies. What we must foresee and what we must do occur in a 3D space surrounding a point of reference corresponding to our selves. If we are to understand the biophysical basis of self, then we must first give a plausible neuronal account of how the human brain is able to represent an object-filled space from an egocentric perspective.

When we look at the world in front of us, we perceive a stable, coherent arrangement of objects and environmental features in a spatially extended layout. But on any given visual fixation, our window of sharp foveal vision allows us to register clearly a region of only 2–5 degrees of the scene before us. Saccadic eye movements present us with a sequence of scattered glimpses of our spatially extended visual environment where all sharply defined visual stimuli are superposed on the fovea. A question then arises: how can the visual system distentangle its fovea-centered images and construct an integrated representation of its surrounding environment, not in a fovea-centered frame, but in an egocentric spatial frame? This problem led me to propose the existence of a post-retinal dynamic buffer in the brain which I named the *retinoid system*. Its putative structural and dynamic properties enable it to register and appropriately integrate disparate foveal stimuli into an egocentric representation of an extended 3D frontal scene, as well as perform many other useful perceptual and higher cognitive functions. (Among these is the representation of our body image in egocentric space, e.g., my head is above my toes.) Neuronal details of the retinoid system have been modeled and tested in computer simulations (Trehub, 1977, 1978, 1991).

3. The retinoid system

The retinoid system registers information in visual space and projects afferents to higher visual centers. It organizes successive retinocentric visual inputs into coherent representations of object layout in 3D space. It also receives input from higher visual centers and can serve as a visual scratch pad with spatially organized patterns of excitation stored as short-term memory. The mechanism of temporary storage is assumed to be in

the form of retinotopically and spatiotopically organized arrays of excitatory autaptic neurons. These are cells which have their own axon collaterals in feedback synapse with their own dendrites or cell body (Lubke, Markram, Frotscher, & Sakmann, 1996; Tamas, Buhl, & Somogyi, 1997; van der Loos & Glaser, 1972). An autaptic cell that receives a transitory suprathreshold stimulus will continue to fire for some period of time if it is properly biased by another source of subthreshold excitatory input. Thus, a sheet of autaptic neurons can represent by its sustained discharge pattern any momentary input pattern for as long as diffuse priming excitation (excitatory bias) is sustained (up to the limit of cell fatigue). If the priming background input is terminated or sufficiently reduced, the discharge pattern which represents the stimulus on the retinoid will rapidly decay (see Trehub, 1991; Fig. 2.5). The problem of registering and combining disparate foveal stimuli into a proper unified representation of a larger real-world scene can be solved by a layered system of interconnected retinoids acting as dynamic postretinal buffers (Trehub, 1977, 1991).

The dynamics of retinoid activity are further enhanced by the neuronal structure shown in Fig. 1. Each retinoid layer is composed of an array of autaptic neurons connected by a balanced grid structure of excitatory and inhibitory interneurons. Shift control cells send axon collaterals in excitatory synapse with selected groups of these interneurons. Any momentary suprathreshold input from an afferent visual array to its homologous autaptic retinoid cells will evoke sustained firing of the retinoid targets if there is a sufficient level of diffuse tonic bias. Thus, retinal stimulation induces a comparable retinoid pattern of spatially organized discharge. At the same time, each autaptic neuron induces a subthreshold, priming excitatory postsynaptic potential (EPSP) in each of its eight contiguous interneurons capable of eliciting both excitatory and inhibitory potentials (IPSP) in their targeted autaptic cells. Any primed interneuron that receives a sufficient increment of excitation from one of the shift control cells will fire and send spike input to its target cell. The retinoid behaves according to the following implicit rules.
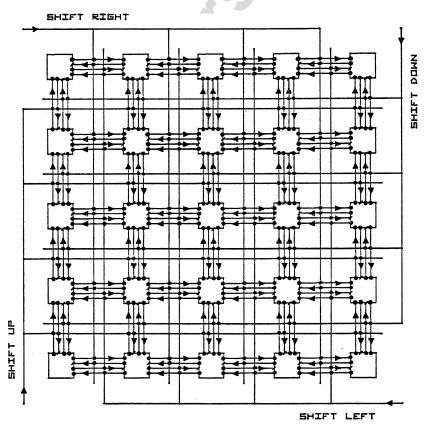

Fig. 1. A retinoid array. Squares represent autaptic cells serving short-term memory. Small filled triangles between autaptic cells represent excitatory and inhibitory interneurons. Dots represent excitatory synapses. Oblique slashes represent inhibitory synapses. Shift-control cells designated by direction of effect.

1. If an autaptic cell that is not discharging (OFF) receives sufficient EPSP from an interneuron, it will fire (turn ON).
2. If an autaptic cell that is ON receives IPSP from an interneuron, it will turn OFF unless it receives simultaneously EPSP from another interneuron, in which case it will remain ON.
3. If diffuse excitatory bias to the retinoid falls below a critical level, all cells in the retinoid turn OFF.

An underlying assumption of the model is that mechanisms in the retina and lower visual centers extract an edge-based transform of a fixated object and project this pattern of retinotopic excitation to the retinoid system. The visual pattern will be captured in short-term memory as a spatial analog on the normally biased retinoid. The captured pattern can then be spatially translated in any direction in retinoid space by appropriate pulses from the shift-control command cells. For example, a spike train from the shift-right line will transfer standing activity (via interneurons) from any active autaptic cell to its adjacent autaptic cell on its right and erase (via interneurons) the activity in the previously active donor (on the left) unless the donor is also receiving transferred excitation from its left-adjacent autaptic cell. By this means, sustained command discharges from the shift-right control cell will move the entire retinoid excitation pattern to the right in successive increments of a single autaptic cell. Similarly, commands to shift left, up, or down will move the captured pattern in the appropriate direction. The higher the frequency of the discharge spikes within the shift-control pulses, the more rapidly will the retinoid representation move. The longer the pulse train is sustained, the greater will be the distance through which the representation is moved. Appropriate control-pulse sequences of shift right/left and shift up/down can move a captured stimulus pattern to any position within retinoid space (Trehub, 1977, 1991).

Patterns of neuronal activity can be translated and positioned within retinoid space either by the execution of control plans in higher cognitive centers which drive shift-control neurons, or by neuronal signals from lower mechanisms. Automatic spatial translation of retinoid patterns can occur when shift-control neurons are driven by voluntary eye and/or head movements, by the discharge of cells in lower-level visual mechanisms which detect the motion of objects in the field of view, and by signals from the vestibular organs. When pattern translation is induced by eye or head motion, the direction in which a representation is shifted on a retinoid sheet corresponds to the direction of gaze; the extent to which it is shifted is proportional to the visual angle between the current fixation and the egocentric reference projection, which I call the *normal foveal axis*. This is the coordinate of origin for the egocentric frame within the retinoid system. The normal foveal axis corresponds to the line of sight of the fovea when the eyes are straight ahead, the head is unturned, and the shoulders are square with the upright body.

3.1. Scene assembly from a 2D retinal projection

Assume a complete frontal scene to be all discriminable objects within an environment subtended by 180 degrees centered on the normal foveal axis. Since the receptive field for sharp foveal vision is 2–5 degrees at most, the eyes and/or the head must pivot to scan the whole frontal scene. A sequence of excitation patterns will be evoked on the foveal region of the retina which must be represented in the brain in a way that conserves their projected spatial relationships in the 2D frontal plane.

Fig. 2 shows a module consisting of a number of retinoid layers (R1–R5). The autaptic cells of the layers are in homologous and reciprocal projection to the corresponding cells of the neighboring layers in the manner shown in this figure. The first layer (R1) receives retinotopic input (Sij) from the foveal and near parafoveal region of the retina. Retinoid R1 is a translation retinoid in which the output of its shift-control neurons is modulated by eye and/or head position. The direction and extent of pattern translation on this layer is directly proportional to the degree of eye and/or head shift from the normal foveal axis. As each scene segment is registered on R1 and appropriately shifted in accordance with eye and head position, it is then immediately transferred to R2 in its proper relative egocentric location and "erased" in R1 by reset inhibition to all cells in R1. With successive fixations, a representation of multiple objects and their layout in the frontal scene can be assembled on R2. This will provide the larger scene context for particular sensory inputs.
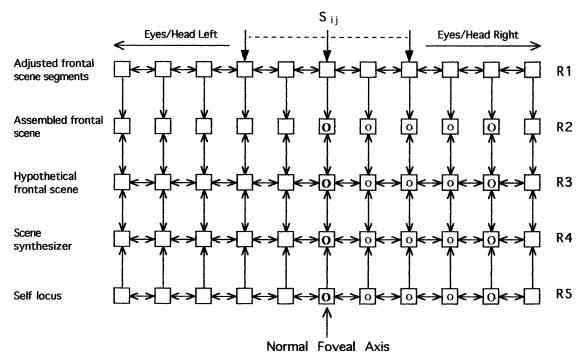
Fig. 2. Stacked retinoid registers (R1–R5). Imagine a dimension of each retinoid array projecting orthogonally to the plane of the page in replicated autaptic cells. Arrows indicate the cells to which excitation can be transferred within the array. Self locus designated by **O**. Heuristic self-locus designated by O. Excursion path designated by o.

3.2. Hypothetical scenes

The retinoid layer labeled R3 in Fig. 2 can construct patterns of autaptic cell excitation representing hypothetical arrangements of visual objects. It is a translation array that can receive veridical information from the assembled frontal scene in R2, as well as internally fabricated patterns projected from the retinoid complex called a scene synthesizer (R4). The scene synthesizer consists of two translation retinoids coupled as a functional unit so that either can serve as a stable storage buffer while the other may be engaged in moving excitation patterns from one place to another in its representational space. The source of the patterns may be output from other retinoids, images recalled and projected from the memory store of a synaptic matrix, or excitation tracings created by movements of the heuristic self-locus (both described below). The combined output of the retinoids in the scene synthesizer can be projected to R3. Thus, whereas R2 in Fig. 2 represents the current veridical scene, by combining and spatially rearranging veridical, remembered and/or fabricated object representations, R3 can assemble complex hypothetical scenes as constructions of our creative imagination.

3.3. Self location

In order to engage in effective behavior within the local environment or plan our future actions, we must have some internal representation of the actual and possible spatial relationships between our self and other significant objects in veridical space. This is accomplished by the *self-locus* retinoid together with the other retinoid registers shown in Fig. 2.

The self-locus retinoid is a translation retinoid with no afferent input but with the capability of point-to-point output projection to other retinoids. It is distinctive in that it maintains a uniquely coded compact region of tonic autaptic discharge with its central "point" located at the center of its retinoid space. This fixed position of the self-locus lies on the normal foveal axis. The autaptic neurons representing the central self-location (the "home" location of the self) are constantly active in the normal waking state, and autaptic discharge originating at the self-locus can be spatially translated to any position on the retinoid surface by the usual shift-command controls. The heuristic locations and excursion paths of self-locus excitation can then be projected to other retinoids and can be combined with real and/or hypothetical objects and scenes represented by

spatiotopic images on their surfaces to construct internal maps of goal regions, obstacles, and direct and indirect paths to goals.

3.4. Selective attention

The ability to move excitation from the source point of the self-locus to selected regions of retinoid space also provides an important means of selective attention. Autaptic cells in regions of a retinoid that are stimulated by the added local excitation of a heuristic self-locus excursion are preferentially primed and marked relative to other cells in retinoid space. Cells in a primed region respond more quickly and vigorously than those in unprimed regions, and the heuristic self-locus marker can serve as a spatial reference so that simply reversing a self-locus excursion command can back-translate any stimulus pattern which might fall on the heuristic coordinate. Back-translation will shift the retinoid representation of the visual environment so that the region of interest falls on the normal foveal axis, the source coordinate of the self-locus (see Fig. 3). This kind of normalization of stimulus patterns on a fixed axis in retinoid space simplifies the formidable problem of object learning and recognition (Trehub, 1991, 1997). It should also be noted that priming and back-translation of selected regions of egocentric space induced by excursions of the heuristic self-locus make it possible to shift visual attention covertly without corresponding eye movements (McMains & Somers, 2004; Posner, 1980; Posner, Snyder, & Davidson, 1980; Shulman, Remington, & McClean, 1979).

3.5. Finding and positioning pattern centroids

Learning, recognition, and the recall of object images from our internal store of previously learned patterns is accomplished by a neuronal mechanism which I call the *synaptic matrix* (Trehub, 1967, 1975, 1983, 1991). The synaptic matrix works interactively with the retinoid system. For efficient learning and subsequent recognition of objects it is desirable that their projected retinal patterns at each fixation be positioned post-retinally
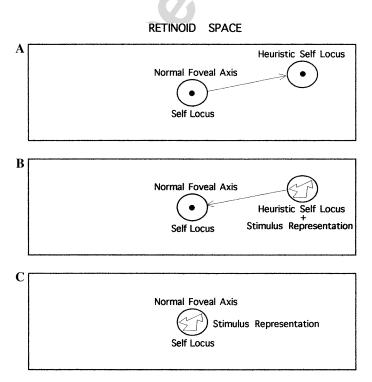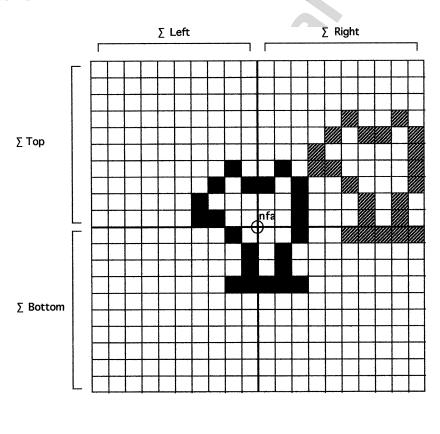
Fig. 3. Selective spatial attention by heuristic excursion of the self locus. (A) A region of retinoid space is primed at the resting location of the heuristic self-locus. (B) A stimulus pattern is evoked in the primed region of the retinoid. Reversal of the initial excursion command translates the stimulus pattern and the heuristic self-locus back to the normal foveal axis. (C) Stimulus is in the standard location on the normal foveal axis where it can be projected to the synaptic matrix to be learned or recognized.

in a standard way on the normal foveal axis before they are gated to the mosaic-cell array, which is the retinotopically organized neuronal input and a part of the imaging circuitry of the synaptic matrix.

But saccadic fixations on salient regions of the visual field do not ensure that the image of an object within a circumscribed region will be centered either on the retina or on the normal foveal axis even if the retinoid representation of that region is translated back to the reference coordinate of the self-locus. One way of accomplishing optimal neuronal positioning for input to the synaptic matrix is to provide a mechanism for finer adjustment which ensures that an object's excitation pattern on a retinoid will be shifted automatically so that its centroid falls as close as possible on the normal foveal axis.

Imagine a retinoid structure organized quadrantally with each quadrant of its surface receiving retinotopic afferents from its respective retinal quadrant (see Fig. 4). On any given fixation, if the excitation of a captured pattern is summed independently over each quadrant of the retinoid and if the relative magnitudes of the summed discharges are used to drive the shift control cells, we then have a neuronal mechanism that can align the centroid of any parafoveal stimulus with the central axis of retinoid space (the normal foveal axis). In the algorithm shown at the bottom of Fig. 4, each autaptic cell represents one unit of excitation. The difference in total excitation between the left and right hemifields is compared to a threshold that represents error tolerance (ET), where error is determined by the degree of mismatch in excitation. If error tolerance is exceeded, a signal is sent to the appropriate shift control cell to drive the retinoid pattern in a direction that reduces error.

$$[\, (\, \Sigma \, \text{Left} - \Sigma \, \text{Right} \,) > \text{ET} \,] ====> \text{Shift Right}$$

$$[\, (\, \Sigma \, \text{Right} - \Sigma \, \text{Left} \,) > \text{ET} \,] ====> \text{Shift Left}$$

$$[\, (\, \Sigma \, \text{Top} - \Sigma \, \text{Bottom} \,) > \text{ET} \,] ====> \text{Shift Down}$$

$$[\, (\, \Sigma \, \text{Bottom} - \Sigma \, \text{Top} \,) > \text{ET} \,] ====> \text{Shift Up}$$

Fig. 4. Quadrantally organized retinoid. Intersection of the vertical and horizontal axes defines the normal foveal axis. Each square represents an autaptic cell. The excitation evoked by a stimulus pattern is summed independently for each of the vertical and horizontal hemifields. The algorithm shown at the bottom of the figure shifts a pattern so that its excitation centroid falls on the normal foveal axis (nfa). Diagonally hatched cells show the initial retinoid location of a parafoveal stimulus. Solid black cells show the stimulus with its centroid on the normal foveal axis after it has been shifted to balance mismatches in quadrantal excitation.

Exactly the same kind of mechanism adjusts the position of the stimulus pattern over the top and bottom fields. In this way, mismatches of excitation across the retinoid hemifields control the activity of shift control cells so that excitation is balanced over the retinoid quadrants (within error tolerance) and brings the centroid of an image to the normal foveal axis (Trehub, 1990, 1991).

3.6. Depth perception and stereoscopic vision

The mechanisms that have been outlined so far deal with visual-cognitive processes in 2D space. Fig. 5 illustrates the arrangement of sensory and neuronal elements in a 3D retinoid that are required for depth
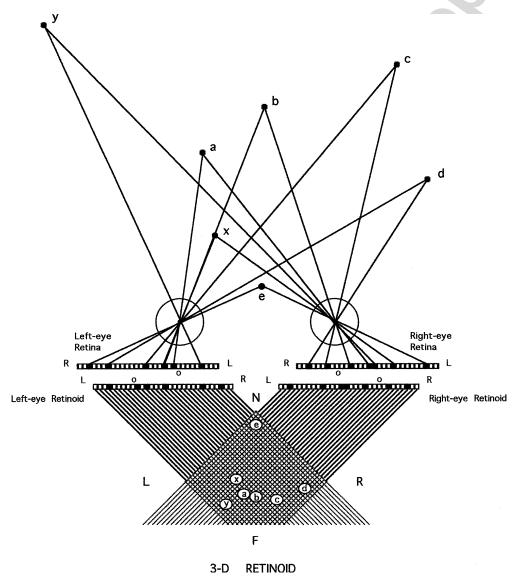

Fig. 5. The 3D retinoid. Imagine all retinoid elements stacked orthogonally to the plane of the page. Lower-case letters at the top of the figure represent objects in space with their projections to the left-eye and right-eye retinas. Line-of-sight projections from each object terminate on the filled squares which represent the stimulated retinal cells. Left-eye and right-eye retinoids are shown just below their corresponding retinas. Filled retinoid squares represent stimulated autaptic cells. L and R indicate left and right visual fields. Notice that the retinal representations are transformed at each retinoid to correspond with the relative egocentric locations of objects in terms of monocular visual angle. Diamond shaped cells in the 3D retinoid represent correlation clusters that are innervated by intersecting retinotopic axon projections from the left-eye and right-eye retinoids. N indicates the near visual field. F indicates the far visual field. Lower-case letters within the 3D retinoid indicate correlation clusters having maximum evoked activity in response to the corresponding objects in the visual field. The N–F dimension defines the Z-axis. The rows of correlation clusters orthogonal to the Z-axis define the Z-planes (depth planes).

perception on one horizontal plane in the binocular visual field. The plane is defined along the *X*-axis (L–R) for the horizontal dimension and along the *Z*-axis in depth. Imagine many identical structures stacked on top of each other to represent the *Y*-axis in 3D perception. Each small square in the strips designated "left-eye retina" and "right-eye retina" is assumed to contain a uniform 2D array of retinal receptors and associated ganglion cells. The retinal image at each eye is projected retinotopically as an edge transform to its associated monocular retinoid, where it is mapped to the egocentric frame and shifted in accordance with the direction and degree of eye deviation from the normal foveal axis. The position of the visual pattern on each monocular retinoid will correspond to the relative angular position of objects and their features in the egocentric frontal field of each eye.

Each small diamond-shaped box in the 3D retinoid shown at the bottom of Fig. 5 contains a single cluster of retinotopically organized autaptic neurons corresponding to the ganglion cells that provide their sensory input through the two monocular retinoids. Thus, every autaptic cell in each small 3D cluster receives two afferent axons, one relayed from its corresponding ganglion cell in the left-eye retinoid and the other from its corresponding cell in the right-eye retinoid. Each diagonal string of cell clusters in the 3D retinoid shown in Fig. 5 represents a line-of-sight array. The principle function of each cluster is to perform a cross-correlation between the micropatterns of excitatory inputs arriving from the left eye and the right eye. Rows of correlation clusters that are orthogonal to the *Z*-axis in the figure define depth planes (also called *Z*-planes). On the basis of binocular disparity, the regions of highest micropattern correlation at intersections of the line-of-sight arrays locate visual features in depth and squelch false targets. This network can integrate monocular retinoid outputs in a binocular system to provide depth perception and stereopsis in 3D space (see Trehub, 1978, 1991, for a detailed account of the correlation mechanism in the cell clusters of the 3D retinoid system as well as a random dot simulation test of the model).

3.7. Size constancy

When an object moves from a location close to an observer to a distant location, its projected image on the retina becomes progressively smaller as its visual angle decreases. As the retinal image decreases in size, the autaptic-cell representations of the object in the monocular retinoids and in the 3D retinoid also shrink in size. However, despite the size change in retinal and retinoid images as a function of object distance, we normally judge a given object to be of constant size (Boring, 1940; Graham, 1951). A specialized arrangement of axonal projections from the autaptic cells in the 3D retinoid to the mechanism for object recognition (the synaptic matrix) provides a biophysical basis for the perception of size constancy.

The model assumes that in the central receptive field there is a structure of connectivity from 3D retinoid cells to the synaptic matrix that roughly normalizes the projected size of an excitation pattern in the 3D retinoid as a function of the particular *Z*-plane that the pattern occupies. Autaptic cells in the nearest 3D *Z*-plane are mapped as input to their corresponding retinotopic connections with the mosaic array in the synaptic matrix. Cells in the more distant *Z*-planes diverge to project to increasingly more eccentric coordinates on the input to the synaptic matrix in accordance with the visual distance they represent. Thus, retinal images that become smaller as a function of increasing object distance are enlarged in compensatory fashion as they are mapped onto the synaptic matrix to be recognized. Notice, however, that the number of cells activated by an object will always decrease in proportion to its distance, with a corresponding decrease in visual resolution. This is another instance of the visual system's multiple representation of the properties of objects in the visual world. The relative retinal size of an object is represented within the 3D retinoid at the same time that its relative intrinsic size is represented in its output to the synaptic matrix, the brain's putative mechanism for learning and object recognition.

Suppose the retinal representation of an object were to remain constant in size, while its representation in the 3D retinoid were to shift to more distant *Z*-planes. If we arrange for this to happen, we should perceive the constant retinal image increasing in size as its retinoid distance increases, because of the expanding projection from the more distant *Z*-planes in the size-constancy mechanism. To experience this size illusion, fixate the center of the black square in Fig. 6 for approximately one minute. If you now fixate the cross below the square, you will see a square image that is brighter than the surrounding surface. Move the page further away and the square after-image will appear to grow larger. Move the page closer than the original fixation distance
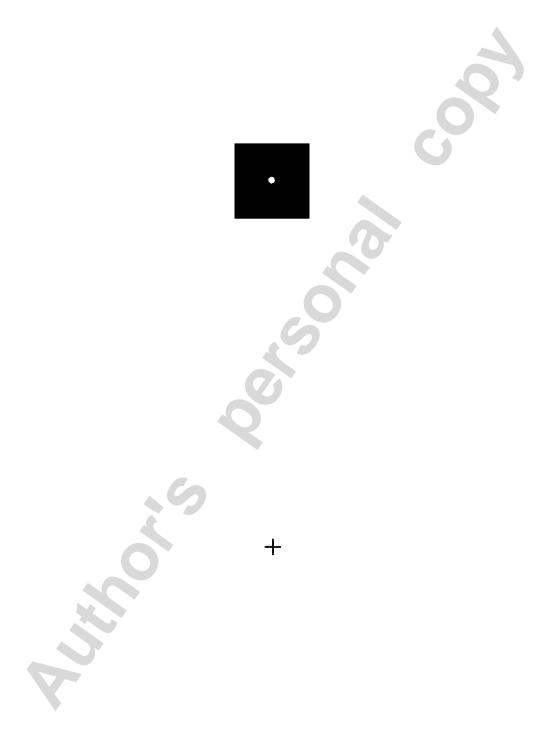
Fig. 6. An inversion of the size-constancy effect in viewing an after image.

and the bright square will appear to become smaller than its original size. It is important to recognize that the retinal after-image remains constant in size during all changes in fixation distance. The phenomenon is predictable from the properties of the retinoid system and its projections to the synaptic matrix. As the page is viewed at different distances, the 3D mechanism shifts the concurrent retinal after-image to the Z-plane that corresponds to the relative distance of the fixated page. In normal perception, the size of an object's retinal image varies inversely with fixation distance and the compensatory projection of the size-constancy mechanism conserves our perception of the object's intrinsic size (size normalization). However, when the retinal after-image of the square is captured at different distances in the 3D retinoid, its retinal-image size does not vary as a function of fixation distance, thus the compensatory retinoid projections produce an inversion of the size-constancy effect.

The size normalization structure plays a very important role in object learning and recognition. Since the normalized pattern of input excitation to the synaptic matrix will be roughly equivalent in size for any given aspect of an object over a wide range of arbitrary distances, a detector cell that is tuned to a particular object viewed at one distance will be an effective detector of that object at other distances. Thus, an object need not be learned separately for each distance at which it might appear in relation to the viewer.

4. Some supporting evidence

Two key assumptions of the retinoid model are: (1) visually induced neuronal excitation patterns can be spatially translated over arrays of spatiotopically organized neurons, and (2) excitation patterns can be held in short-term memory within the retinoids by means of self-synapsing neurons called autaptic cells. I made these assumptions originally because they provided the theoretical grounding for a brain mechanism capable of processing visual images in 3D space very efficiently and because they seemed physiologically plausible (Trehub, 1977, 1978, 1991).

More recent experimental results provide direct neurophysiological evidence supporting these assumptions. Munoz, Guitton, and Pelisson (1991) recorded from single units in the superior colliculus of cats. They found that at the start of an orienting gaze shift, a zone of neuronal discharge was evoked at the collicular locus that coded the desired gaze displacement. As the gaze shift proceeded, the zone of activation moved continuously across the retinotopically coded motor map toward the rostral (foveal) pole of the superior colliculus. Duhamel, Colby, and Goldberg (1991) recorded from visually responsive single units in the lateral intraparietal area in monkeys that were required to perform fixation and saccade tasks in order to receive a reward. They found that by translation of activation, these cells update the retinal coordinates of remembered stimuli to generate a continuously accurate retinotopic representation of egocentric visual space. Employing functional magnetic resonance imaging (fMRI), Liu, Pestilli, and Carrasco (2005) demonstrated with human subjects that if a visual target location is precued, there is an improvement in visual performance that is tightly correlated with a larger fMRI response in corresponding retinotopic areas. Of additional interest was the finding that fMRI enhancement progressively increased from striate to extrastriate areas. With single-unit recording in monkeys, Constantinidis and Steinmetz (2005) found that posterior parietal cortex automatically detects and encodes the spatiotopic location of salient visual stimuli even when they are unrelated to a behavioral task. With the use of optical imaging of intrinsic signals, Raffi and Siegel (2005) were able to show that a monkey's internally generated locus of attention is correlated with an 800–860 μm patchy topological architecture across the cortical surface of the inferior parietal lobe. Using an advanced optical technique employing a calcium-sensitive indicator in vivo, Ohki, Chung, Ch'ng, Kara, and Reid (2005) demonstrated that the organization of functional maps in the visual cortex can have single-cell precision in their resolution of visual input.

The neuroanatomical observations of van der Loos and Glaser (1972) provided support for the proposed existence of autaptic cells in the early retinoid model (Trehub, 1977). Since then, there have been many neurophysiological studies which give direct evidence for the widespread presence of autaptic units in the cortex. For example, Lubke et al. (1996) examined the frequency of autapses (synapses of an axon collateral to the dendrite or body of the same cell) in rat neocortex and found that 80% of all layer-5 pyramidal cells which they analyzed in somatosensory cortex could be classified as autaptic neurons. Tamas et al. (1997) examined the visual cortex of cats and found autaptic cell activity to be common in several different cell types within brain areas 17 and 18.

There is also abundant evidence that pyramidal cells in the visual system and in other connected brain systems show the kind of temporally sustained activity that the retinoid mechanism requires. For example, Fuster and Jervey (1981) recorded single cell activity in the inferotemporal cortex of monkeys during a delayed matching-to-sample task. They found that a substantial number of cells maintained an elevated frequency of spike discharge during the 17 s memory retention period of the task (i.e., when the sample stimulus was turned off). Graziano, Hu, and Gross (1997) recorded from single units in the ventral premotor cortex of monkeys. An object was presented within the visual receptive field of individual neurons and then the lights were turned off and the object was removed. The investigators reported that "A subset of the neurons continued to respond in the dark as if the object were still present and visible". Fried, MacDonald, and Wilson (1997) recorded from single units in the medial temporal lobe of humans who were given a recognition memory test for faces and objects. They reported that "Traces [measured by spike activity] of exposure to faces or objects were found a few seconds after stimulus removal as well as 10 hours later". They also claimed that there were some neurons which maintained a record of previous stimulus presentation that was more accurate than the person's conscious recollection. Application of fMRI in humans who perform spatial working memory tasks reveals an area in the superior frontal sulcus which seems specialized for spatial cognitive processing (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998). Retinoid mechanisms have the structural and dynamic properties well suited to perform such tasks, and it appears from this and the previously cited studies that the retinoid system might be distributed over several cortical regions including the temporal and parietal areas and the superior frontal sulcus.

A particularly striking finding of hemispatial neglect in patients with brain lesions involving the right temporal–parietal–occipital junction (Bisiach & Luzzatti, 1978) also lends support to the retinoid model of spatial representation. In this study, patients were asked to imagine that they were standing in the main square in Milan which was a very familiar setting for them. They were first instructed to imagine themselves facing the cathedral and to describe what they could see in their "mind's eye." They reported a greater number of details to the right than to the left of their imaginary line of sight, often neglecting prominent features on the left side. When they were asked to perform the same imagery task, facing away from the cathedral, they were able to report previously neglected details within the right half of the imaginal perspective but ignored items in the left half that they had reported just a few moments before.

This example of hemispatial neglect indicates that the deficit is not due to a loss of sensory input, a fault in sensory processing, or a selective memory loss. It appears instead to be due to a failure in the internal representation of egocentric space, or in the ability to recognize objects represented within a particular region of egocentric space. In terms of the retinoid model, the deficit can be explained by (a) damage to the left hemifield of the retinoid system (a representational loss), (b) selective damage to the part of the shift-control mechanism that translates retinoid images from left to right in order to bring them to the normal foveal axis for recognition in the synaptic matrix (an analytical loss), or (c) selective damage to the part of the shift-control mechanism that drives the heuristic self-locus into the left hemifield of retinoid space (an attentional loss).

In addition to evidence provided by neurophysiological studies, brain imaging and clinical observations, it can be shown that the retinoid model, which was not designed to account for illusions, can nevertheless explain many visual illusions as natural consequences of its basic neuronal structure and dynamics. These included illusions induced by motion after-effects (e.g., Riggs & Day, 1980), the Müller-Lyer illusion, and the most puzzling of all natural illusions, the moon illusion (about which I will have more to say below). Moreover, a completely novel phenomenon which I have called the pendulum illusion, was successfully predicted on the basis of the detailed properties of the retinoid mechanism (Trehub, 1991). In this illusion, when an object is moved in horizontal back-and-forth motion behind an occluding screen and is viewed through a triangular aperture in the screen, it appears to swing like a pendulum pivoting at the vertex of the aperture.

5. Self as the neuronal origin of retinoid space

With this summary of the neuronal structure and dynamics of the retinoid system as a foundation, we can now take a closer look at the properties of the self locus within the brain's retinoid representation of 3D space. I will argue that the autaptic neurons at the self locus (SL) of the retinoid structure constitute

the core biophysical substrate for our sense of self. I will also argue that heuristic excursions of SL excitation in retinoid space together with synaptic couplings of SL cells with sensory and executive mechanisms, and with neuronal tokens of semantic processes give us the rich diverse contents of our phenomenal experience. In this theoretical formulation, we can think of the SL as the "home" location of the neuronal activity which grounds the self.

The first thing to notice is that the self is located on the normal foveal axis and on the nearest Z-plane within the 3D retinoid (see Fig. 8). This is the "point" of origin in our egocentric space—according to the model, the most intimate location that the human can represent. Heuristic excursions and traces of autaptic-cell activity representing the self emanate from the home position of the self and project excitation paths to selected targets in 3D retinoid space.

In attempting to relate the biophysical embodiment of oneself to one's conscious experience, I think it is important to ask if there is a minimal or irreducible awareness on which the richer content of consciousness depends. My own persistent reflections suggest that whenever I am conscious, there is a minimal level of awareness that can best be described as a sense of centeredness within a surround. There need be no other experiential content—just a sense of immutable location. This must be an inner-determined event because my worldly location ranges widely. But the minimal self stays fixed at the center of all experience while the body travels wherever. The extent to which this sense of fixed location is experienced in the general population is an interesting empirical question which can only be answered by careful investigation of introspective reports. In any event, excitatory activity in a compact cluster of autaptic cells at the origin of the normal foveal axis of the 3D retinoid provides a plausible biophysical substratum for the raw sense of centeredness in space, which I take to be the minimal property of consciousness. Moreover, according to the retinoid model, the number of autaptic cells that are traversed (and successively discharged) in the shortest excursion of the heuristic self-locus from its origin to any other autaptic-cell location (target) is assumed to be the biophysical substratum for a person's phenomenal experience of an object's distance.

5.1. Phenomenal space is not isotropic

It is usually taken for granted that the extent of physical space is uniform in all directions. If our perception of a physical space that is isotropic is determined by the structure of the brain's retinoid system, it would seem reasonable to assume that the dimensional properties of the 3D retinoid are also uniform in all spatial directions and distances. On this assumption, we would expect that our perception of the relative distance to the sky ahead and to the sky above should be the same. But it has long been known that we do not see the sky this way. Even before the writings of Ptolemy, observers reported that the "dome" of the sky appeared to be flattened from horizon to horizon (Smith, 1738). The vertical distance of the sky is perceived to be closer than its horizontal distance. Why should this be?

It is commonly accepted that our brain, and our visual-cognitive system in particular, has been shaped by evolution to cope with the ecologically significant demands of the human environment. Examples of this principle are clearly evident in biobehavioral investigations of the visual system in lower animals. It has been found that visual receptors and their associated neuronal structures in amphibians and birds are allocated in rough accordance with needs related to the survival of the species (Fite, 1976; Ingle, 1968; Spinelli, 1987). On this principle, if the location of objects in particular regions of egocentric space is a matter of little ecological importance, then we might expect a relative reduction in neuronal resources devoted to the representation of those regions, with the savings possibly invested (by evolutionary processes) in more useful brain structures.

Can we characterize regions of human egocentric space as having more or less ecological utility? While we cannot assign precise values to the relative utility of sharply defined extrapersonal regions, there are considerations that persuasively suggest the relative ecological values of broadly specified regions in egocentric space. Space within immediate reach, for example, is critical for a wide variety of behaviors essential to human survival, development and socialization (e.g., grasping, eating, tool use and assembly, social interaction, etc.). Beyond the bounds of this near surround, space within the range of quick natural locomotion may contain objects and features which demand immediate response. As the regional locus of terrestrial space increases in distance from the observer, the objects contained in it tend to diminish in ecological urgency. Thus, the

brain's reduction in distance discrimination as the distance of visual targets increases results in no serious loss of the ability to cope in the natural environment. In addition, when we consider the location of objects that have ecological significance, we find that as egocentric distance increases, the maximum vertical angle of regard naturally decreases. For example, an apple on a tree branch that is within reaching distance may be directly overhead, a vertical sighting of 90 degrees; but any earthbound object at a horizontal distance of 1000 feet would have to be 1000 feet tall at the point of fixation to demand an angle of regard of as much as 45 degrees.

If evolution has conserved neuronal resources for 3D viewing of horizon-limited terrestrial space, why should it not devote even fewer biological resources for the 3D representation of extraterrestrial objects? After all, such objects were, except for the briefest (modern) period of human experience, beyond the scope of possible bodily contact, and demanded only philosophical speculation. I have proposed that there is a structural anisotropy in the brain's 3D retinoid system so that retinoid representations of objects within regions of increasing egocentric distance are progressively collapsed onto nearer Z-planes as the elevation in their lines of sight departs from the plane of the normal horizon (Trehub, 1991). In this way, valuable neuronal tissue is conserved for the adaptive functions more closely related to our survival.

Recall that the vertical distance of the sky is perceived to be closer than its horizontal distance. This phenomenon follows naturally from a vertical/horizontal anisotropy in the neuronal structure of the 3D retinoid which provides our intimate experience of physical space. As we look around us, the locus of our attention moves in accordance with our visual search. In the retinoid model, shifts in attention are real neuronal events—excursions of the heuristic self-locus (focal autaptic-cell discharge) over selected coordinates of retinoid space. When we scan the sky, the neuronal excitation of the heuristic self-locus traces the outer bounds of retinoid space. Thus, the phenomenal distance from ourself to the sky is closer as the line of sight approaches the egocentric vertical axis, and greater as the line of sight approaches the horizontal axis (the horizon).

6. Two decisive phenomena as evidence for the retinoid model

The acceptance or rejection of any scientific theory rests on the weight of empirical and logical evidence for and against the theory. In the history of science, we have occasionally been confronted with phenomena that were explained/predicted by a candidate theory and were so surprising and counter-intuitive that their appearance was taken as critical evidence decisively in support of the theory. A famous example is the observation, in 1919, of a displacement of the light from Mercury as it grazed the eclipsed sun, an effect that was predicted by Einstein's theory of gravitation. Below I describe two phenomena which, I believe, lend decisive support to the retinoid model of brain representation.

6.1. The moon illusion

When we look at the moon at or near the horizon it is seen as much larger than it appears overhead, even though its projected retinal size remains unchanged at a visual angle of approximately 0.5 degrees. This is the moon illusion—perhaps the oldest and most puzzling illusory phenomenon (Hershenson, 1989; Ross & Plug, 2002). The retinoid model was developed to explain how the human brain represents an object-filled 3D space; it was not designed to provide a satisfactory account of the moon illusion, which serves no adaptive purpose. In a serendipitous way, however, it became apparent that the moon illusion will occur as a natural consequence of two basic structural properties of the 3D retinoid (Trehub, 1991 see in particular Figs. 14.6 and 14.7). When the combined effect of viewing distance and vertical angle of regard results in a shift of a retinal stimulus of constant size from a very far Z-plane in the 3D retinoid to a nearer Z-plane, then the axonal projections of the image from the retinoid to the input array of the synaptic matrix, which normally conserves size constancy, will shrink in size. Because of the vertical/horizontal anisotropy in the 3D retinoid, as the moon rises above the horizon, the most distant Z-planes that have a neuronal substrate to represent its increasing elevation correspond to progressively nearer regions of 3D space. Thus, the image projected to the mosaic array of the synaptic matrix is that of a rising moon diminishing in size, and this is what we see.

The moon illusion gives strong support to the retinoid model of spatial representation and object perception. The model also demonstrates an intimate and systematic relationship between a measurable aspect of conscious experience and biophysical details of the putative brain mechanism which, I propose, is the foundation of our normal experience of physical space and its perceived content.

6.2. Seeing-more-than-is-there

If a narrow vertically oriented aperture in an otherwise occluding screen is fixated while a visual pattern is moved back and forth behind it, the entire pattern may be seen even though at any instant only a small fragment of the pattern is exposed within the aperture. This phenomenon of anorthoscopic perception was reported as long ago as 1862 by Zöllner (Zöllner, 1862). More recently, Parks (1965), McCloskey and Watkins (1978), and Shimojo and Richards (1986) have published work on this striking visual effect. McCloskey and Watkins introduced the term *seeing-more-than-is-there* to describe the phenomenon and I have adopted it in abbreviated form as SMTT. The following experiment was based on the SMTT paradigm.

6.2.1. Procedure

1. Subjects sit in front of an opaque screen having a long vertical slit with a very narrow width, as an aperture in the middle of the screen. Directly behind the slit is a computer screen, on which any kind of figure can be displayed and set in motion. A triangular-shaped figure in a contour with a width much longer than its height is displayed on the computer. Subjects fixate the center of the aperture and report that they see two tiny line segments, one above the other on the vertical meridian. This perception corresponds to the actual stimulus falling on the retinas (the veridical optical projection of the state of the world as it appears to the observer).

2. The subject is given a control device which can set the triangle on the computer screen behind the aperture in horizontal reciprocating motion (horizontal oscillation) so that the triangle passes beyond the slit in a sequence of alternating directions. A clockwise turn of the controller increases the frequency of the horizontal oscillation. A counter-clockwise turn of the controller decreases the frequency of the oscillation. The subject starts the hidden triangle in motion and gradually increases its frequency of horizontal oscillation.

Results: As soon as the figure is in motion, subjects report that they see, near the bottom of the slit, a tiny line segment which remains stable, and another line segment in vertical oscillation above it.

As subjects continue to increase the frequency of horizontal oscillation of the almost completely occluded figure there is a profound change in their experience of the visual stimulus.

At an oscillation of ~2 cycles/s (~250 ms/sweep), subjects report that they suddenly see a complete triangle moving horizontally back and forth instead of the vertically oscillating line segment they had previously seen. This perception of a complete triangle in horizontal motion is strikingly different from the line segment oscillating up and down above a fixed line segment which is the real visual stimulus on the retinas.

As subjects increase the frequency of oscillation of the hidden figure, they observe that the length of the base of the perceived triangle decreases while its height remains constant. Using the rate controller, the subject reports that he can enlarge or reduce the base of the triangle he sees, by turning the knob counter-clockwise (slower) or clockwise (faster).

3. The experimenter asks the subject to adjust the base of the perceived triangle so that the length of its base appears equal to its height.

Results: As the experimenter varies the actual height of the hidden triangle, subjects successfully vary its oscillation rate to maintain approximate *base–height equality*, i.e., lowering its rate as its height increases, and increasing its rate as its height decreases.

This experiment demonstrates that the human brain has internal mechanisms that can construct accurate *analog* representations of the external world. Notice that when the hidden figure oscillated at less than 2 cycles/s, the observer experienced an event (the vertically oscillating line segment) that corresponded to the visible event on the plane of the opaque screen. But when the hidden figure oscillated at a rate greater than 2 cycles/s, the observer experienced an internally constructed event (the horizontally oscillating triangle) that corresponded to the almost totally occluded event behind the screen.

The experiment also demonstrates that the human brain has internal mechanisms that can accurately track *relational* properties of the external world in an analog fashion. Notice that the observer was able to maintain an approximately fixed one-to-one ratio of height to width of the perceived triangle as the height of the hidden triangle was independently varied by the experimenter.

These and other empirical findings obtained by this experimental paradigm were predicted by the neuronal structure and dynamics of a putative brain system (the retinoid system) that was originally proposed to explain our basic phenomenal experience and adaptive behavior in 3D egocentric space (Trehub, 1991). It seems to me that these experimental findings provide conclusive evidence that the human brain does indeed construct analog representations of the external world.

7. Interaction between analogical and symbol/token representations

Encompassed within the bounds of egocentric space are all the objects of our immediate experience. Exteroceptive and interoceptive sensory transducers provide parallel arrays of afferent excitation to the higher centers of the brain. Particular afferent patterns are systematically related to the different kinds of physical events for which the individual transducers are adapted. In the case of vision, the retinas send to the higher visual centers a time-varying retinotopic pattern of neuronal excitation that captures the significant two-dimensional properties of an object. When the same retinal information is processed in the retinoid system, the pattern of autaptic-cell excitation in the 3D retinoid captures some of the three-dimensional properties of the object and its location in 3D egocentric space. These excitation patterns in the brain can be considered analogical representations or images of the visual objects that induce the patterns. This is so because the brain patterns conserve retinotopy and are essentially similar in topological structure to the objects that they represent.

In the normal course of cognitive processing, many of the brain's analogical representations of environmental objects are grouped together into separate categories to which we can apply symbols or common names, e.g., dog, cat, house, etc. When I see a dog, its analogical representation in my brain (S) can selectively and reliably evoke the specific neuronal activity (T) that will trigger my utterance "dog". We can think of such a sequence as a neuronal event which warrants the proposition that S (neuronal analog) is a proper brain referent of T (neuronal symbol/token), or $S \rightarrow T$. Moreover, if someone asks me to think of a dog or if I read the word "dog", I am likely to imagine the visual appearance of a dog. This implies that the neuronal event (T') which is triggered by the word "dog" (spoken or printed) can evoke an analogical representation of a dog (S') in my brain. We can think of this sequence as a neuronal event which warrants the proposition that T' (neuronal symbol/token) is a proper token for S' (neuronal analog), or $T' \rightarrow S'$. Taken together, this suggests that in the human brain there are selective reciprocal evocations between analogical representations and symbolic/token representations. Thus, in the general case, $S \leftrightarrow T$ where S is the referent of T, and T is the token of S.

In the proposed model of the cognitive brain, specialized neurons that are selectively (synaptically) linked to the brain's representation of any culturally accepted word/symbol and to the neuronal routine for its production, must discharge in order to trigger the mechanisms that enable us to process, utter, or write any common word or formal symbol. These neurons, which are essential components of synaptic matrices, are called class cells (Ω). They are our basic biological symbols/tokens for all learned objects, events, images, and concepts. In effect, they give us the private brain "names" needed to index our experience independently of any culture-bound lexicon. Crucially, in addition to providing the biological machinery necessary to categorize the world, neuronal properties of the synaptic matrix also enable the discharge of a class-cell token (Ω) to evoke a neuronal analog (image) of the object that is tokened (Trehub, 1977, 1979, 1991). In this way, $S \leftrightarrow T$ is realized in the synaptic matrices of the human brain, and the referential foundation of our semantics is established. Fig. 7 shows a block-flow diagram of the activation sequence from visual stimulation through retinoid representation, pattern recognition, lexical assignment, semantic processing, and back to the lexical referents (the induced images in the mosaic array of synaptic-matrix 1).

8. The theater of consciousness

Much of the important enriching content in our sense of space is provided by reciprocal evocation among visual images and tokens embodied in the recurrent loop of the synaptic matrix and its coupling with the
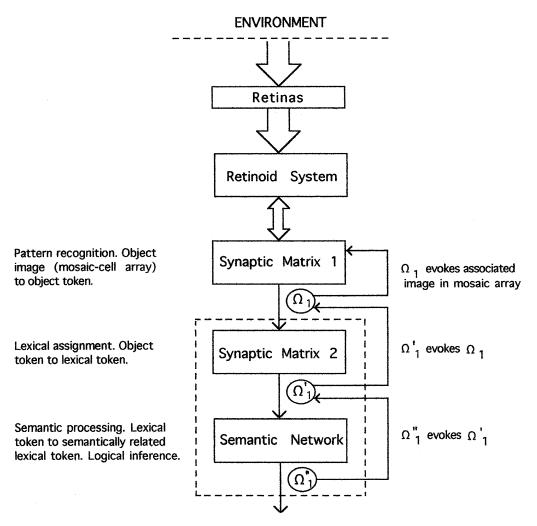
ENVIRONMENT

Pattern recognition. Object image (mosaic-cell array) to object token.

Lexical assignment. Object token to lexical token.

Semantic processing. Lexical token to semantically related lexical token. Logical inference.

Retinas

Retinoid System

Synaptic Matrix 1

Ω_1

Ω_1 evokes associated image in mosaic array

Synaptic Matrix 2

Ω'_1

Ω'_1 evokes Ω_1

Semantic Network

Ω''_1

Ω''_1 evokes Ω'_1

Fig. 7. Forward and backward chaining of Ω tokens. Adaptive mapping binds lexical tokens to appropriate object tokens and to their sensory images.

retinoid system. But more to the point of this paper is the notion that if the self-locus itself were innately tokened, and if this token of one's self-locus were synaptically linked to the web of all of one's other learned neuronal tokens and images, then this master token, with its immediate referent at the very center of ego space, would be the nexus and the cohering agent of all personal experience. In *The Cognitive Brain* (Trehub, 1991), I proposed that such a token exists in the human brain, and I named it the I-token (symbolized as **I**!). Wherever I may be, **I**! is always right here. In principle, it is possible that you and I can occupy the same spatial location at different times, and you and I can even be at approximately the same place at the same time. But you and **I**! not only can never occupy the same space, we are in totally incommensurable spaces. I take this fact to be the essence of one's sense of individuality of self. To forestall any misunderstanding, I should emphasize that the self is not precariously tied to the fate of a single **I**! cell. It is instead assumed that, distributed in some part of the human brain, there are a number of equivalent I-tokens that are activated in parallel and have redundant connectivity, as do other kinds of neuronal tokens. Thus, the self is normally a robust unitary system. It is interesting to consider that if, for any reason, mutually inhibitory synaptic pathways should develop among some of the I-tokens that are all normally activated in parallel, then sets of competitive **I**!s would organize mutually exclusive selves in a single individual. One might conjecture that this is the neuronal basis for reported cases of multiple personality.

What I have described so far is fundamentally important for our experience of self, but it covers only a limited aspect of the range of tokened brain activity that is synaptically linked to **I**!. Another aspect of

cognition needs to be mentioned here in order to address an essential characteristic of any individual in the context of interpersonal and intrapersonal communication.

8.1. Belief

Humans are famously able to use sentences to communicate with others and with themselves (e.g., I must remember to buy milk today). I have proposed that the biophysical basis for this faculty lies in a particular variant of the synaptic matrix—a semantic network which gives us the capacity to form simple sentential propositions, make logical inferences, and compose narratives (Trehub, 1991). A crucially important aspect of each personality is his/her beliefs. We think of our beliefs as an intimate part of us—defining attributes of our self. Among all the propositions we may think or utter, those we *believe* have a special status and are most likely to influence our behavior. Our beliefs determine what we sincerely assert and how we credit the assertions of others. Problem: if sentential propositions are embodied in the neuronal machinery of the brain, and if belief is attached only to some propositions and not to others, how does the brain represent which propositions are truly its beliefs?

The lexical strings which make up our sentences can be generated by syntactically controlled neuronal-token strings in a semantic network. Because the brain must discriminate between propositions that are believed and those that are treated as subjunctive or false, I have suggested that the brain marks belief by linking a unique neuronal token which signifies belief to those privileged propositions that pass a test of personal validity. And because personal belief is such an intimate part of us, I have proposed that only neuronal propositions that are synaptically linked to the I-token and therefore accompanied by the parallel discharge of **I!** are true expressions of the brain states of belief (Trehub, 1991). For example:

⟨ I can read English ⟩**I!** ⟨ This is true ⟩**I!** ⟨ This is false ⟩
⟨ I can read Sanskrit ⟩ ⟨ This is true ⟩ ⟨ This is false ⟩**I!**
⟨ The human mind is a product of evolution ⟩**I!**

where each lexical item designates its neuronal token.

Notice the distinction between the neuronal token ⟨I⟩ which is synaptically coupled to the word "I" and the neuronal token **I!** which is synaptically coupled to the self locus and to a vast web of other tokened events in the brain.

Like other tokens that are reciprocally linked to semantically related lexical tokens, the discharge of **I!** can evoke its attached propositions in the semantic network of belief. So if you were asked what you believe, you would be able to produce a long list of items "off the top of your head"—"I believe that I am talking to you. I believe that ...", etc. And, just as in the case of other kinds of assertions, many of the lexical tokens, their phonological expressions, and their referential analogs would be represented somewhere within the egocentric space of your 3D retinoid. Details of the neuronal mechanisms that locate these representations within retinoid space are described in Trehub (1991, pp. 129–130, Figs. 7.11 and 7.12).

8.2. The 3D retinoid as the stage of consciousness

The prevailing opinion in the study of mind is that there is no one place in the brain where consciousness all comes together. Dennett (1991) has presented an extended argument in support of the view that the brain creates multiple drafts of experience which somehow constitute what we call consciousness. I think this is true as far as it goes, but a more scientifically useful model of the cognitive brain should offer a more detailed and informative explanation of our conscious experience. It is important to recognize that consciousness is not an all-or-none phenomenon. Consciousness has at least a few distinct and different aspects which I think allow us to get a better understanding of its nature.

As a first approximation within the present theoretical framework, we can distinguish at least three different levels of consciousness. The first (C1) is minimal awareness—what I have referred to as the primitive sense of centeredness within a surround. The brain correlate of C1 would be excitatory activity, beyond some critical

threshold, in the cluster of neurons that constitute the self locus, and in the other autaptic cells of the 3D ret-inoid. It is assumed that this would normally occur only if there is sufficient corticipetal arousal from the retic-ular activating system (Steriade, 1996). The second level would, in addition to C1, include activation of **I!** so that consciousness C2 would be characterized by a minimal sense of oneself at the origin of egocentric space. At the third level (C3), alert awareness, the conscious brain would include activation of all the sensory/cog-nitive mechanisms that are synaptically linked to **I!**. The principal contributors to conscious experience at C3 are depicted in Fig. 8.

It is often claimed that the common metaphorical reference to consciousness as a theater in which we are the observer of events on a stage, is based on a mere Cartesian illusion. But illusions should not be dis-missed so lightly. They are real events caused by real biophysical processes in the specialized machinery of the human brain (Trehub, 1991). As I have shown earlier, there is considerable evidence for the claim that our phenomenal space is strictly limited and determined by the structural and dynamic properties of the 3D retinoid, which is an egocentric topological analog of our natural behavioral space. I suggest that all images and tokens which we experience as having denotable spatial locations are really neuronally pro-jected, by mechanisms outside the retinoid system, to the regions of 3D retinoid space corresponding to our phenomenal experience. According to this view, the notion of a stage seems a reasonable metaphor to describe an important function of the 3D retinoid. But if the retinoid system provides our phenomenal stage, who or what is the audience?

8.3. Evaluative mechanisms as audience

The answer, although a bit complicated, points to a credible biophysical grounding for our normal first-per-son experience. Surely there is no homunculus in our head observing excitation patterns in retinoid space.
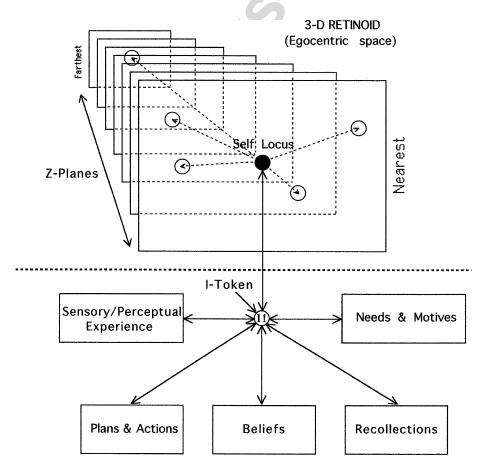

Fig. 8. The self system. The self locus anchors the I-token (**I!**) to the retinoid origin of egocentric space. **I!** has reciprocal synaptic links to the neuronal tokens of all sensory/cognitive processes.

In one sense, what goes on in retinoids is just a part of the many multiple neuronal representations of our experience. But, I maintain, activity in this system is a coherent manifold of all our experience in our phenomenal present. The retinoid system is a complex piece of neuronal machinery which provides a 3D perceptual buffer and workspace (Baars, 1988, 1997). It presents patterns of excitation that other neuronal mechanisms can decompose, analyze, categorize, and evaluate for a multitude of adaptive purposes. On metaphorical grounds, I do not think it stretches matters much to think of specialized neuronal mechanisms such as synaptic matrices, semantic, and affective/hedonic circuits (see Trehub, 1991, pp. 153–168) which categorize/evaluate the cellular activity presented on the retinoid "stage" of C3 as something like a critical observing audience. Moreover, in the normal brain, because of the egocentric perspective of the retinoid system and the binding properties of **I!**, the outputs of all evaluative brain mechanisms (all members of the "audience") are implicit aspects of a unitary self in C3. In addition, recall that in order to make any particular object or setting in retinoid space available to the synaptic matrix for subsequent evaluation, a "spotlight" of attention embodied in a selective shift of the heuristic self-locus must focus on the object or region of our immediate interest. This makes the self an active rather than a passive participant in the "theater of consciousness". Herein is the foundation for our strong sense of personal agency.

What happens on the stage of retinoid space is never a finished production. Metaphorically, it is always a work in progress with constantly changing scenes and cast, directed partly by the current environment and partly by the sensory and evaluative neuronal mechanisms in the "dark" of the theater as they respond to the brain's changing motives and tactics in its encounters with the natural world. There are, however, two constant aspects to the production on the stage of consciousness—one is the space-time envelope of the retinoid stage itself, and the other is the ubiquitous participating self. So, in a metaphorical sense, the retinoid system is the "illuminated" stage of consciousness, the evaluative/executive neuronal mechanisms (the audience) are off-stage. The other supporting adaptive processes of the brain and its body might be loosely thought of as the stage hands. I resist the temptation to propose a counterpart to the metaphorical playwrite.

9. Conclusion

This theoretical model builds on the assumption that our experiences of space, self, belief, and all that we can possibly say about consciousness are absolutely limited by the biophysical structure and dynamics of their embodiment in each cognitive brain. I have given an overview of some key mechanisms which, I hypothesize, account for these phenomena.

To sum up: the retinoid system embodies our raw conscious experience of egocentric space with the self as its neuronal origin. The I-token enables the brain to evoke current, retrospective, and prospective neuronal stages of our sensible self as situated within the egocentric retinoid space of our sensible world. It also provides a neuronal flag of personal belief, enriching self identity and attaching a biological marker to those internal propositions which we must treat most seriously. In order to have a self, we must have an I-token with intact synaptic coupling to the self-locus neurons in the retinoid system. If our synaptic connections to some of the mechanisms shown below the dashed line in Fig. 8 were damaged, we would suffer cognitive impairment but we would still have an **I!**—conserving our sense of a unitary self. But if the supporting excitatory link with the self locus were interrupted, **I!** could not serve its proper functional role and we would no longer experience our self as a localized entity centered in our phenomenal world.

As a final thought, I believe it is worth noting that it was possible to draw this causal account of conscious content only after the development of a detailed large-scale theoretical model of neuronal mechanisms and systems underlying human cognition (Trehub, 1991), in particular the specification of the retinoid system.

References

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge University Press.

Baars, B. J. (1997). *In the theater of consciousness*. Oxford University Press.

Bisiach, E., & Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex, 14*, 129–133.

Constantinidis, C., & Steinmetz, M. A. (2005). Posterior parietal cortex automatically encodes the location of salient stimuli. *Journal of Neuroscience, 25*, 233–238.

Boring, E. G. (1940). Size constancy and Emmert's law. *American Journal of Psychology, 53*, 293–295.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, G. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science, 279*, 1347–1351.

Dennett, D. C. (1991). *Consciousness Explained.* Little Brown and Company.

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1991). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science, 255*, 90–92.

Fite, K. V. (1976). *The amphibian visual system.* Academic Press.

Fried, I., MacDonald, K. A., & Wilson, C. (1997). Single unit activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron, 18*, 753–765.

Fuster, J. M., & Jervey, J. P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science, 212*, 952–955.

Graham, C. H. (1951). Visual Perception. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology.* Wiley.

Graziano, M. S. A., Hu, X. T., & Gross, C. G. (1997). Coding the location of objects in the dark. *Science, 277*, 239–241.

Hershenson, M. (1989). ed. The Moon Illusion. Erlbaum.

Ingle, D. (1968). Visual releasers of prey-catching behavior in frogs and toads. *Brain, Behavior, and Evolution, 1*, 500–518.

Liu, T., Pestilli, F., & Carrasco, M. (2005). Transient attention enhances perceptual performance and fMRI response in human visual cortex. *Neuron, 45*, 469–477.

Lubke, J., Markram, H., Frotscher, M., & Sakmann, B. (1996). Frequency and dendritic distribution of autapses established by layer 5 pyramidal neurons in the developing rat neocortex: Comparison with synaptic innervation of adjacent neurons of the same class. *Journal of Neuroscience, 16*, 3209–3218.

McCloskey, M., & Watkins, J. W. (1978). The seeing-more-than-is-there-phenomenon: Implications for the locus of iconic storage. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 533–564.

McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron, 42*, 677–686.

Munoz, D. P., Guitton, D., & Pelisson, D. (1991). Control of orienting gaze shifts by the tectoreticulospinal system in the head-free cat. III. Spatiotemporal characteristics of phasic motor discharges. *Journal of Neurophysiology, 66*, 1642–1666.

Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P., & Reid, R. C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature, 433*, 597–603.

Parks, T. (1965). Post-retinal visual storage. *American Journal of Psychology, 78*, 145–147.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174.

Raffi, M., & Siegel, R. M. (2005). Functional architecture of spatial attention in the parietal cortex of the behaving monkey. *Journal of Neuroscience, 25*, 5171–5186.

Riggs, L. A., & Day, R. H. (1980). Visual aftereffects derived from inspection of orthogonally moving patterns. *Science, 208*, 416–418.

Ross, H., & Plug, C. (2002). *The Mystery of the Moon Illusion.* Oxford University Press.

Shimojo, S., & Richards, W. (1986). Seeing shapes that are almost totally occluded: A new look at Park's camel. *Perception and Psychophysics, 39*, 418–426.

Shulman, R. L., Remington, R. W., & McClean, J. P. (1979). Moving attention through visual space. *Journal of Experimenta Psychology: Human Perception and Performance, 5*, 522–526.

Smith, R. (1738). *A Compleat System of Opticks* (Vol. 1). England: Cambridge.

Spinelli, D. N. (1987). A Trace of Memory: An Evolutionary Perspective on the Visual System. In M. A. Arbib & A. R. Hanson (Eds.), *Vision. Brain, and Cooperative Computation.* MIT Press.

Steriade, M. (1996). Arousal: Revisiting the reticular activating system. *Science, 272*, 225–226.

Strawson, G. (1997). The self. *Journal of Consciousness Studies, 4*, 405–428.

Tamas, G., Buhl, E. H., & Somogyi, P. (1997). Massive autaptic self-innervation of GABAergic neurons in cat visual cortex. *Journal of Neuroscience, 17*, 6352–6364.

Trehub, A. (1967). Learning Machines and Methods. Patent 3,331,054 (Washington, DC: U.S. Patent Office).

Trehub, A. (1975). Adaptive pattern processing in the visual system. *International Journal of Man-Machine Studies, 7*, 439–446.

Trehub, A. (1977). Neuronal models for cognitive processes: Networks for learning, perception and imagination. *Journal of Theoretical Biology, 65*, 141–169.

Trehub, A. (1978). Neuronal model for stereoscopic vision. *Journal of Theoretical Biology, 71*, 479–486.

Trehub, A. (1979). Associative sequential recall in a synaptic matrix. *Journal of Theoretical Biology, 81*, 569–576.

Trehub, A. (1983). Neuronal model for episodic learning and temporal routing of memory. *Cognition and Brain Theory, 6*, 483–497.

Trehub, A. (1990). A confusion matrix for hand-printed alphabetic characters: testing a neuronal model. In M. L. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Computational and clinical approaches to pattern recognition and concept formation.* Erlbaum.

Trehub, A. (1991). *The cognitive brain.* MIT Press.

Trehub, A. (1997). Sparse coding of faces in a neuronal model: interpreting cell population response in object recognition. In J. W. Donahoe & V. P. Dorsal (Eds.), *Neural network models of cognition: biobehavioral foundations.* Elsevier Science.

van der Loos, H., & Glaser, E. M. (1972). Autapses in neocortex cerebri: synapses between a pyramidal cell's axon and its own dendrites. *Brain Research, 48*, 355–360.

Zöllner, F. (1862). Uber einer neue Art anorthoscopischer Zerrbilder. *Annalen der Physik und Chemie, 27*, 477–484.