

Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition*

ANTONIO R. DAMASIO

University of Iowa College of Medicine

Abstract

Damasio, A.R., 1989. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33: 25–62.

This article outlines a theoretical framework for the understanding of the neural basis of memory and consciousness, at systems level. It proposes an architecture constituted by: (1) neuron ensembles located in multiple and separate regions of primary and first-order sensory association cortices (“early cortices”) and motor cortices; they contain representations of feature fragments inscribed as patterns of activity originally engaged by perceptuomotor interactions; (2) neuron ensembles located downstream from the former throughout single modality cortices (local convergence zones); they inscribe amodal records of the combinatorial arrangement of feature fragments that occurred synchronously during the experience of entities or events in sector (1); (3) neuron ensembles located downstream from the former throughout higher-order association cortices (non-local convergence zones), which inscribe amodal records of the synchronous combinatorial arrangements of local convergence zones during the experience of entities and events in sector (1); (4) feed-forward and feedback projections interlocking reciprocally the neuron ensembles in (1) with those in (2) according to a many-to-one (feed-forward) and one-to-many (feedback) principle. I propose that (a) recall of entities and events occurs when the neuron ensembles in (1) are activated in time-locked fashion; (b) the synchronous activations are directed from convergence zones in (2) and (3); and (c) the process of reactivation is triggered from firing in convergence zones

*This work was supported by NINCDS grant PO1 NS19632. I thank my associates Hanna Damasio, Gary Van Hoesen, and Daniel Tranel for helping me shape many of the ideas summarized here, over the past decade. I also thank other colleagues who read previous versions of this manuscript over the past few years and made numerous helpful suggestions: Patricia Churchland, Victoria Fromkin, Jack Fromkin, Edward Klima, Francis Crick, Terry Sejnowski, Jaques Paillard, Marge Livingstone, David Hubel, Freda Newcombe, Ursula Bellugi, Arthur Benton, Peter Eimas and Albert Galaburda. Requests for reprints should be sent to Antonio R. Damasio, Professor and Head, Department of Neurology, University of Iowa Hospitals and Clinics, Iowa City, IA 52242, U.S.A.

and mediated by feedback projections. This proposal rejects a single anatomical site for the integration of memory and motor processes and a single store for the meaning of entities of events. Meaning is reached by time-locked multi-regional retroactivation of widespread fragment records. Only the latter records can become contents of consciousness.

Introduction

This proposal describes a neural architecture capable of supporting the experiences that are conjured up in recall and are used for recognition, at the level of systems that integrate macroscopic functional regions. It arose out of dissatisfaction with available accounts of the neural basis of higher behaviors, especially those implicit in center localizationism, behaviorism, functional equipotentiality, and disconnection syndrome theory.

The title captures the two principal notions in the proposal. First, perceptual experience depends on neural activity in multiple regions activated simultaneously, rather than in a single region where experiential integration would occur. Second, during free recall or recall generated by perception in a recognition task, the multiple region activity necessary for experience occurs near the sensory portals and motor output sites of the system rather than at the end of an integrative processing cascade removed from inputs and outputs. Hence the term retroactivation to indicate that recall of experiences depends on reactivation close to input and output sites rather than away from them.

The two critical structures in the proposed architecture are the fragment record of feature-based sensory or motor activity, and the convergence zone, an amodal record of the combinatorial arrangements that bound the fragment records as they occurred in experience. There are convergence zones of different orders; for example, those that bind features into entities, and those that bind entities into events or sets of events, but all register combinations of components in terms of coincidence or sequence, in space and time. Convergence zones are an attempt to provide an answer to the binding problem, which I see as a central issue in cognitive processing, at all taxonomic levels and scales of operation.

The adult organization described here operates on the basis of neurobiological and reality constraints. During interactions between the perceiver's brain and its surround, those constraints lead to a process of feature, entity, and event grouping based on physical structure similarity, spatial placement, temporal sequence, and temporal coincidence. The records of those perceptuomotor interactions, both at fragment level and at combinatorial level, are inscribed in superimposed and overlapped fashion; yet, because

of the different conditions according to which they are grouped, they become committed to separate neural regions. In cognitive terms I will refer to these processes as domain formation (a creation of relatively separable areas of knowledge for faces, man-made objects, music, numbers, words, social events, disease states, and so on), and recording of contextual complexity (a recording of the temporal and spatial interaction of entities within sets of concurrent events). In neural terms I will refer to these grouping processes as regionalization.

The same type of neuron ensembles, operating on the same principles, constitutes the substrate for different cognitive operations, depending on the location of the ensemble within the system and the connections that feed into the ensemble and that feed back out of it. Location and communication lines determine the topic of the neuron ensemble. The connectivity of functional regions defines the systems-level code for cognitive processes.

The neuroanatomical substrates for this organization are:

- (1) primary and early association cortices, both sensory and motor, which constitute the substrate for feature-based records;
- (2) association cortices of different orders, both sensory and motor, some limbic structures (entorhinal cortex, hippocampus, amygdala, cingulate cortices), and the neostriatum/cerebellum, which constitute the substrate for convergence zones;
- (3) feed-forward and feedback connectivity interrelating (1) and (2), at multiple hierarchical levels, with reciprocal patterns;
- (4) non-specific thalamic nuclei, hypothalamus, basal forebrain, and brainstem nuclei.

The cognitive/neural architecture outlined above can perform: (1) perceptuomotor interactions with the brain's surround; (2) learning of those interactions at the representational level defined above; (3) internal activation of experience-replicative representations in a recall (perception-independent) mode; (4) problem solving, decision making, planning, and creativity; and (5) communication with the environment. All those functions are predicated on a key operation: the attempted reconstitution of learned perceptuomotor interactions in the form of internal recall and motor performance. Attempted perceptuomotor reconstitution is achieved by time-locked retroactivation of fragmentary records, in multiple cortical regions as a result of feedback activity from convergence zones. The success of this operation depends on attention, which is defined as a critical level of activity in each of the activated regions, below which consciousness cannot occur.

According to this proposal, there is no single site for the integration of sensory and motor processes. The experience of spatial integration is brought

about by time-locked multiple occurrences. I thus propose a recursive, iterative design to substitute for the traditional unidirectional processing cascades.

Although the notion of representation covers all the inscriptions related to an entity or event, that is, both fragment and binding code records, the proposal posits that only the multiregional retroactivations of the fragment components become a content of consciousness. The perceptuomotor reconstitutions that form the substrate of consciousness thus occur in an anatomically restricted sector of the cerebrum, albeit in a distributed, multiple-site manner.

In this proposal, and unlike traditional neurological models, there is no localizable single store for the meaning of a given entity within a cortical region. Rather, meaning is reached by widespread multiregional activation of fragmentary records pertinent to a stimulus, wherever such records may be stored within a large array of sensory and motor structures, according to a combinatorial arrangement specific to the entity. A display of the meaning of an entity does not exist in permanent fashion. It is recreated for each new instantiation. The same stimulus does not produce the same evocations at every instantiation, though many of the same or similar sets of records will be evoked in relation to the same or comparable stimuli. The records that pertain to a given entity are distributed in the telencephalon both in the sense that they are inscribed over sizable synaptic populations and in the sense that they are to be found in multiple loci of cerebral cortex and subcortical nuclei.

The proposal permits the reinterpretation of the main types of higher cognitive disorder – the agnosias, the amnesias, and the aphasias – and prompts testable hypotheses for further investigation of those disorders. It also provides a basis for neural hypotheses regarding psychiatric conditions such as sociopathy, phobias and schizophrenia. Several predictions based on this proposal are now being tested in humans, with or without focal brain lesions, using advanced imaging methods and cognitive probes. Some anatomical and physiological aspects of the proposal can be investigated in experimental animals. The concept of convergence zone can be explored with computational techniques.

The need for temporo-spatial integration and its traditional solution

Current knowledge from neuroanatomy and neurophysiology of the primate nervous system indicates unequivocally that any entity or event that we normally perceive through multiple sensory modalities must engage geographically separate sensory modality structures of the central nervous system. Since virtually every conceivable perception of an entity or event also calls for a

motor interaction on the part of the perceiver and must include the concomitant perception of the perceiver's somatic state, it is obvious that perception of external reality and the attempt to record it are a multiple-site neurophysiological affair. This notion is reinforced by the discovery, over the past decade, of a multiplicity of subsidiary functional regions that show some relative dedication not just to a global sensory modality or motor performance but also to featural and dimensional aspects of stimuli (see Damasio, 1985a; Van Essen & Maunsell, 1983; Livingstone & Hubel, 1988, for a pertinent review). The evidence from psychological studies in humans is equally compelling in suggesting featural fragmentation of perceptual processes (see Barlow, 1981; Julesz, 1971; Posner, 1980; Triesman & Gelade, 1980). Early geographic parcellation of stimulus properties has thus grown rather than receded, and the condition faced by sensory and motor representations of the brain's surround is a fragmentation of the inscription of the physical structures that constitute reality, at virtually every scale. The physical structure of an entity (external, such as an object, or internal, such as a specific somatic state) must be recorded in terms of separate constituent ingredients, each of which is a result of secondary mappings at a lower physical scale. And the fragmentation that obtains for concrete entities is even more marked for abstract entities and events, considering that abstract entities correspond to criterion-governed conjunctions of dimensions and features present in concrete entities, and that events are an interplay of entities.

The experience of reality, however, both in ongoing perception as well as in recall, is not parcellated at all. The normal experience we have of entities and events is coherent and "in-register", both spatially and temporally. Features are bound in entities, and entities are bound in events. How the brain achieves such a remarkable integration starting with the fragments that it has to work with is a critical question. I call it the *binding problem* (I use the term binding in a broader sense than it has been used by Treisman and others, to denote the requisite integration of components at all levels and scales, not only in perception but also in recall). The brain must have devices capable of promoting the integration of fragmentary components of neural activity, in some sort of ensemble pattern that matches the structures of entities, events, and relationships thereof. The solution, implicitly or overtly, has been, for decades, that the components provided by different sensory portals are projected together in so-called multimodal cortices in which, presumably, a representation of integrated reality is achieved. According to this intuitively reasonable view, perception operates on the basis of a unidirectional cascade of processors, which provides, step by step, a refinement of the extraction of signals, first in unimodal streams and later in a sort of multimedia and multitrack apparatus where integration occurs. The general

direction of the cascade is caudo-rostral, in cortical terms, and the integrative cortices are presumed to be in the anterior temporal and anterior frontal regions. Penfield's findings in epileptics undergoing electrical stimulation of temporal cortex seemed to support this traditional view (Penfield & Jasper, 1954), as did influential models of the neural substrates of cognition in the post-war period, such as Geschwind's (1965) and Luria's (1966). The major discoveries of neurophysiology and neuroanatomy over the past two decades have also seemed compatible with it. On the face of it, anatomical projections do radiate from primary sensory cortices and do create multiple-stage sequences toward structures in the hippocampus and prefrontal cortices (Jones & Powell, 1970; Nauta, 1971; Pandya & Kuypers, 1969; Van Hoesen, 1982). Moreover, without a doubt, single-cell neurophysiology does suggest that, the farther away neurons are from the primary sensory cortices, the more they have progressively larger receptive fields and less unimodal responsivity (see Desimone & Ungerleider, 1988, for a review and restatement of the traditional view). Until recently, the exception to this dominant view of anterior cerebral structures as the culmination of the processing cascade was to be found in Crick's (1984) hypothesis for a neural mechanism underlying attention.

The purpose of this text is to question the validity of the conventional solution. I doubt that there is a unidirectional cascade. I also question the information-processing metaphor implicit in the solution, that is, the notion that finer representations emerge by progressive extraction of features, and that they flow caudo-rostrally. Specifically, we believe that by using this view of brain organization and function the experimental neuropsychological findings in patients with agnosia and amnesia become unmanageably paradoxical. I also suggest that there is a lack of neuroanatomical support for some requirements of the traditional view, and that there are neuroanatomical findings to support an alternative model. Finally, I believe that available neurophysiological data can be interpreted to support the alternative theory I propose.

Paradoxes and contradictions for the traditional solution

Objections from human studies with the lesion method

If temporal and frontal integrative cortices were to be the substrate for the integration of neural activity on the basis of which perceptual experience and its attempted recall unfold, the following should be found:

(a) That the bilateral destruction of those cortices should preclude the perception of reality as a coherent multimodal experience and reduce experi-

ence to disjointed, modality-specific tracks of sensory or motor processing to the extent permitted by the single modality association cortices;

(b) That the bilateral destruction of the integrative cortices should reduce the quality of even such modality-specific processing, that is, reduce the richness and detail of perception and recall commensurate with the quality obtainable by the level of non-integrative stations left intact;

(c) That the bilateral damage to the rostral integrative cortices should disable memory for any form of past integrated experience and interfere with all levels and types of memory, including memory for specific entities and events, even those that constitute the perceiver's autobiography, memory for non-unique entities and events, and memory for relationships among features, entities, and events.

The results of bilateral destruction of the anterior temporal lobes, either in the medial sector alone or the entire anterior temporal region, as well as bilateral destruction of prefrontal cortices, either in separate sectors or in combination, deny all but a fraction of one of these predictions.

Evidence from anterior temporal cortex damage

It is *not* true that coherent, multimodal, perceptual experience is disturbed by bilateral lesions of the temporal integrative units, and it is *not* true that those lesions cause the perceptual quality of experience to diminish. On the contrary, all available evidence indicates that at both consciously reportable and non-conscious covert levels, the quality of perceptual experience of subjects who have sustained major selective damage to anterior temporal cortices is comparable to controls (see Corkin, 1984; Damasio et al., 1985a,b, 1987). Such subjects can report on what they see, hear, and touch, in ways that observers cannot distinguish from what they themselves see, hear, and touch. A variety of covert knowledge paradigms (e.g., forced recognition and passive skin conductance) indicates that they can also discriminate stimuli, probably on the basis of non-conscious activation of detailed knowledge about the items under scrutiny (Bauer, 1984; Tranel & Damasio, 1985, 1987, 1988). More importantly, the knowledge that such subjects can evoke consciously, at a non-autobiographical level, indicates that ample memory stores of "integrated experience" remain intact after damage to the alleged integrative units. These facts support the contentions: (1) that a considerable amount of integration must take place early on in the system well before higher-order cortices are reached; (2) that integrated information can be recorded there without the agency of rostral integrative units; and (3) that it can be re-evoked there too, without the intervention of rostral integrative structures.

The only accurate prediction regarding the role of alleged integrative units applies to anterior temporal cortices and concerns the loss of the ability to recall unique combinations of representations that were conjoined in experience within a specific time lapse and space unit. That ability is indeed lost, along with the possibility of creating records for new and unique experiences. This is exemplified by the neuropsychological profile of the patient Boswell, whose cerebral damage entirely destroyed, bilaterally, both hippocampal systems (including the entorhinal cortex, the hippocampal formation, and the amygdala), the cortices in anterolateral and anteroinferior temporal lobes (including areas 38, 20, 21, anterior sector of 22, and part of 37), the entire basal forebrain region bilaterally (including the septal nuclei, the nucleus accumbens and the substantia innominata, which contains a large sector of the nucleus basalis of Meynert), and the most posterior part of the orbitofrontal cortices. Boswell's perception in all modalities but the olfactory is flawless and the descriptions he produces of complex visual or auditory entities and events are indistinguishable from those of his examiners. All aspects of his motor performance are perfect. His use of grammar, his phonemic and phonetic processing, and his prosody are intact. His memory for most entities is preserved, and at generic/categorical levels his defect only becomes evident when subordinate specificity is required for the recognition of uniqueness or for the disambiguation of extremely similar exemplars. For instance, he recognizes virtually any man-made object such as a vehicle, tool, utensil, article of furniture or clothing, but cannot decide whether he has previously encountered the specific exemplar, or whether or not it is his. Although he can recognize the face of a friend as a human face, or his house as a house, and provide detailed descriptions of the features that compose them, he is unable to conjure up any event of which the unique face or house was a part, and which belong to his autobiography. In short, his essential perceptuomotor interaction with the environment remains normal provided uniqueness of recognition, recall, or action are not required. Recognition, recall, and imagery operate as they should for large sectors of knowledge at the generic/categorical level.

Evidence from anterior frontal lobe damage

Damage to bilateral prefrontal cortices, especially those in the orbitofrontal sector, is compatible with normal perceptual processes and even with normal memory for entities and events, except when they pertain to complex domains such as social knowledge (Damasio & Tranel, 1988; Eslinger & Damasio, 1985). Bilateral lesions in superior mesial and in dorsolateral cortices cause defects in drive for action, attention, and problem solving, that may secondar-

ily influence perceptual tasks. However, even extensive ablation of virtually the entire prefrontal cortices is compatible with normal perception. The study of Brickner's patient A, of Hebb's and Ackerly and Benton's patients (see Damasio, 1985b for a review), and of our subject EVR (see Eslinger & Damasio, 1985) provides powerful evidence in this regard. Frontal lobe structures, with their multiple loci for the anchoring of processing cascades (Goldman-Rakic, 1988), are even less likely candidates to be the single, global site of integration than their temporal counterparts.

Evidence from damage in single-modality cortices

Perhaps the most paradoxical aspect of these data, when interpreted in light of the traditional view, is that damage in certain sectors of sensory association cortices does affect the quality of some aspects of perception within the sensory modality of those cortices. For instance, damage in early visual association cortices can disrupt perception of color, texture, stereopsis, and spatial placement of the physical components of a stimulus. The range of loss depends on which precise region of visual cortex is most affected (Damasio, 1985a).

The perceptual defect is accompanied by an impairment of recall and recognition. For instance, achromatopsia (loss of color perception) also precludes imaging color in recall (Damasio, 1985; Farah, 1989 and unpublished observations), that is, no other cortices, and certainly no other higher-order, integrative cortices, are capable of supporting the recall of the perceptually impaired feature. The coupling of perceptual and recall impairments is strong evidence that the same cortices support perception and recall. This finding, based on lesion method studies, is in line with evidence from normal human experiments (Kosslyn, 1980). It also suggests an economical approach to brain mapping of knowledge that might obviate the problem of combinatorial explosion faced by the traditional view. In my proposal, the brain would not re-inscribe features downstream from where it perceives them. Furthermore, damage within some sectors of modal association cortices can disturb recall and recognition of stimuli presented through that modality, even when basic perceptual processing is not compromised. The domain of stimuli, and the taxonomic level of the disturbance, depend on the specification of the lesion in terms of site, size, and uni- or bilaterality (Damasio & Tranel, 1989; see also work on category-related recognition defects reviewed in Damasio, 1989; and McCarthy & Warrington, 1988). Lesions within visual association cortices may impair the recognition of the unique identity of faces, while allowing for the recognition of facial expressions, non-unique objects, and visuo-verbal material. Or lesions may compromise object recognition and leave face recog-

nition untouched (Feinberg, Rothi, & Heilman, 1986; Newcombe & Ratcliff, 1974), or compromise reading but not object or face recognition (Damasio & Damasio, 1983; Geschwind & Fusillo, 1966). The key point is that damage in a caudal and modal association cortex *can disrupt recall and recognition at even the most subordinate taxonomic level*. It can preclude the kind of integrated experience usually attributed to the rostral cortices, that is, an evocation made up of multiple featural components, based on different modalities, constituting entities and events. This can happen without disrupting perception within the affected modality and without compromising recall or recognition in other modalities. Damage in modal cortices also disrupts learning of new entities and events presented through the modality (Damasio et al., 1989a).

These findings indicate that a substantial amount of perceptual integration takes place within single-modality cortices, and that knowledge recalled at categoric levels (also known as semantic, or generic)¹ is largely dependent on records and interactions among posterior sensory cortices and the interconnected motor cortices.

It also indicates that recall and recognition of knowledge at the level of unique entities or events (also known as episodic)¹ requires *both* anterior *and* posterior sensory cortices, an indication that a more complex network is needed for intricate subordinate-level mappings and that anterior integrative structures alone are not sufficient to record and reconstruct knowledge at such levels.

The implications are:

(a) that the posterior sensory cortices are sites where fragment records are inscribed and reactivated, according to appropriate combinatorial arrangements (by fragments I mean “parts of entities”, at a multiplicity of scales, most notably at the feature level, for example color, movement, texture, and shape); such cortices are also capable of binding features into entities and thus re-enact the perceptual experience of entities and their operations (“local” or “entity” binding). But posterior cortices cannot map non-local contextual complexity at event level, which is to say they cannot map the spatial and temporal relationships assumed by entities within the multiple concurrent events that usually characterize complex interactions with the environment.

(b) the inscription of contextual complexity, that is, the complexity of the

¹The terms semantic and episodic were proposed by Tulving (1972). Our term generic is largely equivalent to semantic and categorial. Elsewhere in the text I refer generic or categorial knowledge as “supraordinate” or “basic object level” knowledge, and to episodic knowledge as “subordinate level” knowledge. The latter terms are drawn from Rosch’s nomenclature for taxonomic levels (Rosch et al., 1976).

combinatorial arrangement exhibited by many concurrent events (non-local or event binding), requires anterior cortices, although its re-enactment also depends on posterior cortices.

The posterior cortices contain all the fragments with which experiences can potentially be reconstituted, given the appropriate combinatorial arrangement (binding). But as far as combinatorial arrangements are concerned, the posterior cortices contain primarily the records for “local” entity or simple event binding. They do not contain records for “non-local” concurrent event binding and are thus unable to reconstitute experiences based on the contextually complex, multi-event situations that characterize one’s autobiography.

The anterior cortices do contain such non-local, concurrent event binding records. The critical point is that since posterior cortices contain *both* fragment and local binding records, they are essential for *all* experience-replicative operations. Anterior cortices are only required to assist experiences that depend on high-level contextual complexity.

I would predict, based on the above hypotheses, that simultaneous damage in strategic regions of *several* single-modality cortices, for example visual, auditory, somatosensory, in spite of intactness of the so-called rostral integrative cortices, would preclude recognition and recall of a sweeping range of stimuli defined by features and dimensions from those modalities, *both* at generic and episodic levels. The central premise behind my proposal, then, is that extensive damage in “early” sensory cortices is the only way of producing the effect normally posited for destruction of the anterior units, namely the suspension of multimodal recognition and recall, from which would follow the abolition of experiences.

A testable hypothesis drawn from this premise is that damage in intermediate cortices (cortices in parts of areas 37, 36, 35, and 39 that constitute virtual “choke points” for the feed-forward–feedback projections that interlock earlier and higher-order cortices) should have a comparable disrupting effect. There is preliminary evidence that this is so from findings on patients with lesions in these areas (Damasio et al., unpublished; Horenstein, Chamberlin, & Conomy, 1967), and a study is currently under way to analyze additional evidence.

Neuroanatomical and neurophysiological evidence

Leaving aside the fact that no bilateral lesion in a presumed “anterior integrative cortex” is capable of precluding coherent perception of any entity or event, or categorical recall, one might turn around and pose a purely neuroanatomical question: which area or set of areas could possibly function

as a fully encompassing and single convergence region, based on what is currently known about neural connectivity? The simple answer is: none. The entorhinal cortex and the adjacent hippocampal system (hippocampal formation and amygdala) do receive connections from all sensory cortices, and come closest to the mark. Prefrontal cortices, inasmuch as one can envisage their connectivity from neuroanatomical studies in non-human primates, do not fit the bill either. They have no single point of anatomical convergence equivalent to the entorhinal cortex, only separate convergence points with different and narrower admixtures of innervation. The hypothesis suggested by these facts is that the integration of sensory and motor activity necessary for coherent perception and recall must occur in multiple sites and at multiple levels. A single convergence site is nowhere to be found.

In fact, developments in neuroanatomy and neurophysiology have emphasized the notion of segregation while beginning to reveal different possibilities for integration. For instance, Hubel and Livingstone (1987) and Livingstone and Hubel (1984) have demonstrated that separate cellular channels within area 17 are differently dedicated to the processing of color, form and motion. Beyond area 17 the evidence shows:

(1) Early channel separation and divergence into several functional regions revealed by neurophysiological studies (Allman, Miezin, & McGuinness, 1985; Hubel & Livingstone, 1987; Livingstone & Hubel, 1984; Van Essen & Maunsell, 1983), and characterized in part by studies of connectivity (Gilbert, 1983; Livingstone & Hubel, 1987a; Lund, Hendrickson, Ogren, & Tobin, 1981; Rockland & Pandya, 1979, 1981). This form of organization is describable by the attributes divergent, one-to-many, parallel, and sequential.

(2) The existence of back-projections to the feeding cortical origin, capable of affecting processing in a retroactive manner, and capable of cross-projecting to regions of the same level (Van Essen, 1985; Zeki, 1987, personal communication). This anatomical pattern opens the possibility for various forms of local integration.

(3) Existence of convergence into functional regions downstream (projections from visual, auditory, and somatosensory cortices) can be encountered in combinations from two and three modalities, in progressively more rostral brain regions such as areas 37, 36, 35, 38, 20 and 21 (Jones & Powell, 1970; Seltzer & Pandya, 1976, 1978; Pandya & Yeterian, 1985),² a design feature

²The human areas 37 (mesially), 36, and 35 largely correspond to fields TF and TH in the monkey, and to fields TF and TH of von Economo and Koskinas in the human. They are extremely developed in the human, especially area 37. Area 38 corresponds to TG; areas 20 and 21 to TE. Area 39 (the angular gyrus) also represents a major human development and may correspond to expansion of cortices in both posterior superior temporal sulcus and inferior parietal lobule. Area 40 (the supramarginal gyrus) is largely a new human area.

describable by the attributes convergent, many-to-few, parallel, and sequential. In humans, judging from evidence in non-human primates, trimodal combinations are likely to occur in functional regions within Brodmann's areas 37, 36, 35, 38, 39; bimodal combinations are likely in areas 40, 20 and 21.

(4) Existence of further feedback from the latter cortices, that is, "convergence regions", have the power to back-project divergently to the feeding cortices.

The pattern of forward convergence and retrodivergence is repeated in the rostral cortices of the entorhinal and prefrontal regions. For instance, neuron ensembles in higher-order cortices project into the circumscribed clusters found in layer II and superficial parts of layer III of the entorhinal cortex (Van Hoesen, 1982; Van Hoesen & Pandya, 1975a,b; Van Hoesen, Pandya, & Butters, 1975). I describe this design feature as convergent, and few-to-fewer. Convergence continues into the hippocampal formation proper, by means of perforant pathway projections to the dentate gyrus and of projections from there into CA3 and CA1. Convergence is again followed by divergent feedbacks via several anatomical routes: (1) a direct route, using the subiculum and layer IV of the entorhinal cortex, diverges into the cortices that provide the last station of input into the hippocampus (Kosel, Van Hoesen, & Rosene, 1982; Rosene & Van Hoesen, 1977); as noted above, those cortices project back to the previous feeding station; (2) an indirect route, so far only revealed in rodents but possibly present in primates, which feeds back into virtually all previous stations, divergently and in saltatory fashion, rather than in recapitulatory manner (Swanson & Kohler, 1986); (3) an even less direct and specific route, which uses pathways in the fornix and exerts influence over thalamic, hypothalamic, basal forebrain, and frontal structures, all of which in turn, directly and indirectly, can influence the operation of the cerebral cortices in widespread fashion. The latter route provides the cortex with regionally selective or widespread neurochemical influence (e.g., acetylcholine, norepinephrine, dopamine, and serotonin) based on the activity of neurotransmitter nuclei in basal forebrain and brainstem (Lewis et al., 1986; Mesulam, Mufson, Levey, & Wainer, 1983).

The findings clearly indicate that the hippocampus-bound projection systems point as much forward as backward. Furthermore, the convergence noted anteriorly is always partial, never encompassing the full range of sensory and motor processes that may be involved in complex experiences. Precisely the same argument could be presented for the multiplicity of prefrontal cortices that serve as end-points for projections from parietal and temporal regions. The feed-forward projections remain segregated among parallel

streams and are reciprocated by powerful feedbacks to their originating cortices or their vicinity (Goldman-Rakic, 1988).

The fact that the receptive fields of neurons increase dramatically in a caudal–rostral direction has implicitly supported the notion of rostral integration. A look at this issue in the visual system reveals that the size of the receptive field of neurons in area 17 (V_1) is extremely small; it enlarges by as much as one hundred times at the level of V_4 , and at the level of the higher-order cortices of areas 20 and 21 virtually encompasses the entire visual scene (Desimone, Schein, Moran, & Ungerleider, 1985). This gradual enlargement of receptive fields, all the way from small and lateralized to large and bilateral, has been viewed as an indication that anteriorly placed neurons not only see more of the world but represent a finer picture of it (Desimone & Ungerleider, 1989, Perrett et al., 1987). However, nothing in those data indicates that the fewer and fewer neurons that are linked to larger and larger receptive fields contain any concrete representation whatsoever of the perceptual detail upstream or that those neurons are committed and the end-point of multiple-channel processing. Those data are certainly compatible with the proposal I present below: (a) that fewer and fewer neurons placed anteriorly in the system are projected on by structures upstream and thus subtend a broader compass of feed-forwarding regions; (b) that they serve as pivots for reciprocating feedback projections rather than as the recipients and accumulators of all the knowledge inscribed at earlier levels; and (c) that in such a capacity they are intermediaries in a continuous process that systematically returns to early cortices.

The unavoidable conclusion is that, while it is possible to conceive of the integration of sensory processes within a few neuronal regions necessary to define a single entity, it is apparent that no single area in the human brain receives projections from all the regions involved in the processing of an event. More importantly, it is inconceivable that any single region of the brain might integrate spatially all the fragments of sensory and motor activity necessary to define a set of unique events. An answer to this puzzle, namely the ability to generate an integrated experience in the absence of any means to bring the experience's components together in a single spatial meeting ground, might be a trick of timing. It would allow the perceiver or recaller to experience spatial integration and continuity in relation to sets of activity that are spatially discontinuous but do occur in the same time window, an illusory intuition.

A different solution

Following on the evidence and reflections outlined above and incorporating additional neuropsychological and neuroanatomical data, I propose the following solution:

(a) The neural activity that embodies physical structure representations entity occurs in *fragmented fashion and in geographically separate cortices* located in modal sensory cortices. The so-called integrative, rostral cortices of the anterior temporal and prefrontal regions cannot possibly contain such fragmentary inscriptions.

(b) The integration of multiple aspects of reality, external as well as internal, in perceptual or recalled experiences, both within each modality and across modalities, depends on the time-locked co-activation of geographically separate sites of neural activity within sensory and motor cortices, rather than on a neural transfer and integration of different representations towards rostral integration sites. The conscious experience of those co-activations depends on their simultaneous, but temporary, enhancement (here called co-attention), against the background activity on which other activations are being played back.

(c) The representations of physical structure components of entities are recorded in precisely the same neural ensembles in which corresponding activity occurred during perception, but the combinatorial arrangements (binding codes) which describe their pertinent linkages in entities and events (their spatial and temporal coincidences) are stored in separate neural ensembles called *convergence zones*. The former and the latter neuron ensembles are interlocked by reciprocal projections.

(d) The concerted reactivation of physical structure fragments, on which recall of experiences depends, requires the firing of convergence zones and the concomitant firing of the feedback projections arising from them.

(e) Convergence zones bind neural activity patterns corresponding to topographically organized fragment descriptions of physical structure, which were pertinently associated in previous experience on the basis of similarity, spatial placement, temporal sequence, temporal coincidence, or any combination of the above. Convergence zones are located throughout the telencephalon, at multiple neural levels, in association cortices of different orders, limbic cortices, subcortical limbic nuclei, and non-limbic subcortical nuclei such as the basal ganglia.

(f) The geographic location of convergence zones varies among individuals but is not random. It is constrained by the subject matter of the recorded material (its domain), by degree of contextual complexity in events (the

number of component entities that interact in an event and the relations they adopt), and by the anatomical design of the system.

(g) The representations inscribed in the above architecture, both those that preserve topographic/topologic relationships, and those that code for temporal coincidences, are committed to populations of neuron ensembles and their synapses, in distributed form.

(h) the co-occurrence of activities at multiple sites, which is necessary for temporary conjunctions, is achieved by iteration across time phases.

Thus I propose not a single direction of processing, along single or multiple channels, but rather a recursive and iterative form of processing. Such processing is parallel and, because of the many time phases involved in multiple steps, it is also sequential. Convergence zones provide integration, and, although the convergence zones that realize the more encompassing integration are more rostrally placed, the activities that all levels of convergence zone end up promoting, and on the basis of which representations are reconstituted and evoked, actually take place in caudal rather than rostral cortices. And because convergence zones return the chain of processing to earlier cortices where the chain can start again towards another convergence zone, there is no need to postulate an ultimate integration area. In other words, this model can accommodate the astonishing segregation of processing streams that the work of Livingstone and Hubel has revealed so dramatically.

The sensory and motor cortices are thus seen as the distributed and yet restricted sector of the brain on which both perception and recall play themselves out, and on which self-consciousness must necessarily be based. Perception and self-consciousness are assigned the same brain spaces at the border between the world within and the world without.

In the following section I present a framework based on these views and discuss its structures, systems, organization, and operation.

Timelocked multiregional retroactivation: framework, structures, systems organization, and operation

Framework

Because of its origin in mutually constraining sets of cognitive and neural data, the theory developed here is both cognitive and neural. The cognitive architecture implicit in the theory assumes representations that can be described as psychological phenomena and interrelated according to combinatorial semantics and syntax. The proposed neural organization, however, is not

a mere hardware implementation apparatus for any potential type of cognitive processes, in that its specifications severely restrict the range of representations and algorithms that it can implement; that is, it is not likely to implement representations other than the ones its anatomy and physiology embody and are destined to operate. The key level of neural architecture is that of systems of macroscopic functional regions in cerebral cortex and gray matter nuclei.

The theory describes an adult neural/cognitive organization presumed to be relatively stable and yet modifiable by experience, to produce temporary or long-lasting partial reorganizations. The issues of neural and cognitive development are not addressed, nor does the theory deal with microneural specifications at synaptic and molecular levels. However, it does assume that any inscription of perceptuomotor activity is based on a distributed transformation of physiological parameters, occurring over ensembles of neurons at the level of their synapses, according to some variant of Hebbian principles. The theory operates on the basis of neurobiological and reality constraints.

Neurobiological constraints

These correspond to the structural design of the nervous system prior to interactions with the environment: the basic circuitry of cellular structures and their interconnectivity, which can be changed by epigenetic interactions. The design includes neuroanatomically embodied values of the organism (e.g., goals and drives of the species), external and internal spatial reference maps, and a variety of processing biases that are likely to guide, in part, the mapping of interactions with the environment, that is, the domains of knowledge that the brain prefers to acquire and the choice of neural sites to support such knowledge. The effect of these constraints is to provide a certain degree of innate modularization of “faculties” upon exposure to the reality constraints discussed below.

Reality constraints: the world without and the world within

The description of the characteristics of the universe surrounding the brain, both inside and outside the organism, can be made at the multiple levels that current knowledge of philosophy, psychology, physics, chemistry, and biology permit. From my point of view, however, it is sensible to focus the description on the levels from which we derive psychological meaning: (1) a broad range of objects to which I will refer to as entities and which encompass both natural and man-made kinds; (2) the features and dimensions that compose those entities; and (3) the interplay of entities in unique events or

episodes occurring in temporal and spatial units. Thus, the set of reality constraints corresponds to:

(1) The existence of concrete entities external to both brain and organism, and external to the brain but internal to the organism (somatic). External entities are themselves composed of various aggregated features and dimensions in an entity-intrinsic space (the space defined by the physical limits of the entity) and are, in turn, placed within an entity-extrinsic space (the coordinate space where the entity and other entities lie or move). Internal entities consist of: (a) motor interactions of the organism with external entities by means of movements in hands, head, eyes, and whole body; (b) baseline somatic states of internal milieu and of smooth and striated musculature during interaction with external entities; and (c) modification of somatic states triggered by and occurring during interaction with external entities.

(2) The existence of abstract entities are criterion-governed conjunctions of features and dimensions present in the concrete entities outlined above.

(3) The fact that entities necessarily occur in unique interactive combinations called events, and that events often take place concurrently, in complex sets.

Entities are definable by the number of components, the modality range of those components (e.g., single or multiple modality), the mode of assembly, the size of the class formed on the basis of physical structure similarity, their operation and function, their frequency of occurrence, and their value to the perceiver.

As is the case with entities, events can be both external and internal, and both concrete and abstract. The concurrence of many events which characterize regular life episodes generates "contextual complexity", which can be defined by the number of entities and by the relational links they assume as they interplay in such complex sets of events. Naturally, during the unfolding of events, other entities and events are recalled from autobiographical records. The records co-activated in that process add further to the contextual complexity of the experiences that occur within a given time unit. It is thus contextual complexity which sets entities and events apart and which confers greater or lesser uniqueness to those entities and events. In other words, contextual complexity sets the taxonomic level of events and entities along a continuum that ranges from unique (most subordinate) to non-unique (less subordinate and more supraordinate).

Domain formation and recording of contextual complexity

During interactions between the perceiver's brain and its surround, the two sets of constraints lead to some critical operations that can be described as follows from a psychological standpoint:

- (1) domain formation, which is a process of feature, entity, and event grouping based on physical structure similarity, spatial placement, temporal sequence, and temporal coincidence;
- (2) the creation of records of contextual complexity that register the temporal coincidence of entities and their interrelationships within sets of events.

It is on the basis of this psychological-level description and on the evidence that category-related recognition defects can be associated to damage in specific brain loci that we hypothesize neural substrates for different knowledge domains and levels of knowledge processing. It must be noted that for the purposes of modeling we are here inverting the natural order of things: domains exist *because* of neurobiological and reality constraints, not the other way around.

Functional regionalization

The process of regionalization occurs for both fragments of perceptuomotor activity and convergence zones. I conceive it as a way of recruiting a neuron population for a limited range of cortical inputs (and, by extension, to the domain or level defined by the feed-forwarding neuronal populations). In other words, certain topics (at feature, entity, or event level) are assigned to a circumscribed neuronal population. Within that population, however, different synaptic patterns define individual features, or entities, or events. In simple terms one might say that generally similar material stacks up together within the same regions and systems.

As I will discuss further on, the superimposed, overlapped nature of the records poses problems for their appropriate separation during recall. The solution I envisage, and that may appear counterintuitive at first glance, resides with the wealth and complexity of the record at the synaptic level. The greater the number of defining sub-components and distinctive links, the greater the chance of establishing uniqueness at the time of recording and at the time of reactivation.

The key to regionalization is the detection, by populations of neurons, of coincident or sequential spatial and temporal patterns of activity in the input neuron populations. Precisely the same type of neuron ensembles, operating

on precisely the same principles, will constitute the substrate for different cognitive operations depending on the location of the ensemble within the system and the connections that feed into the ensemble and that feed back out of it. In other words, location and communication lines determine the topic of the synaptic patterns within a given neuron ensemble (the domain of a convergence zone), without there being a need to posit special neuron types or special physiological codes in order for convergence zones to serve different domains or cognitive operations.

The nature of representations

Human experiences as they occur ephemerally in *perception* are the result of multiple sensory and motor processing of a collection of features and dimensions in external and internal entities. Specifically they are based on the cerebral representation of concrete external entities, internal entities, abstract entities, and events.

Such representations are interrelated by combinatorial arrangements so that their internal activation in recall and the order with which they are attended, permits them to unfold in a “sentential” manner. Such “sentences” embody semantic and syntactic principles.

In my view, the words of any language are also concrete external entities. The combinatorial semantics and syntax of thought and language might be embodied in the relationships that describe the constitution of entities and events (although the universal grammar behind language may be based on additional language-specific principles and rules).

This cognitive/neural architecture implies a high degree of sharing and embedding of representations. Both the representation of abstract entities and of events are derived from the representation of concrete entities and are thus individualized on the basis of combinatorial arrangement rather than remapping of constituents. The representation of concrete entities themselves share subrepresentations of component features so that individuality is again conferred by combinatorial formulas.

Human experiences, as they occur ephemerally in *recall*, are based on records of the multiple-site and multiple-level neural activities previously engaged by perception. Recalled experiences constitute an attempted reconstruction of perceptual experience based on activity in a set of pertinent sensory and motor cortices, controlled by a reactivation mechanism specified below.

The components of representations

Feature-based fragments

I propose that the experienceable (conscious) component of representations results from an attempt at reconstituting feature-based, topographic or topologically organized fragments of sensory and motor activity; that is, only the feature-based components of a representation assembled in a specific pattern can become a content of consciousness. The maximal size of the feature-based fragment is a critical issue. Stimuli such as human faces, verbal lexical entities, and body parts of the self, must be permanently represented by large-scale fragments on the basis of which rapid reconstitution can occur. It is unlikely that such stimuli would depend on a reconstruction from the smallest-scale level of neural activity (equivalent, for the visual system, to Bela Julesz' textons, 1981). But many fragments are small-scale and can be shared by numerous entities and used interchangeably in the reconstitution attempt.

Convergence zones

1. The structure and role of convergence zones

Because feature-based fragments are recorded and reactivated in sensory and motor cortices, the reconstitution of an entity or event so that it resembles the original experience depends on the recording of the combinatorial arrangement that conjoined the fragments in perceptual or recalled experience. The record of each unique combinatorial arrangement is the binding code, and it is based on a device I call the convergence zone.

Convergence zones exist as synaptic patterns within multi-layered neuron ensembles in association cortices, and satisfy the following conditions: (1) they have been convergently projected upon by multiple cortical regions according to a connectional principle that might be described as many-to-one; (2) they can reciprocate feed-forward projections with feedback projection (one-to-many); (3) they have additional, interlocking feed-forward/feedback relations with other cortical and subcortical neuron ensembles. The signals brought to convergence zones by cortico-cortical feed-forward projections, represent temporal coincidences (co-occurrence) or temporal sequences of activation in the feeding cortices (rather than re-representations of inscriptions contained in the feeding cortices). I envision the binding code as a synaptic pattern of activity such that when one of the projections which feed-forward to it is reactivated, firing in the convergence zone leads to simultaneous firing in all or most of the feedback projections which reciprocated the

feed-forward from the original set. By means of those reciprocating feedback lines, convergence zones can trigger simultaneous activity in all or part of the originally feeding cortices, in a retroactive and divergent manner, according to certain principles of operation specified below. The proposal does not address the issue of the number or size of convergence zones, although it assumes that the zone's size is defined during development as a result of input-output connection patterns, and the patterns of lateral interaction that help structure the ensemble as a unit.

Convergence zones are *amodal*, in that they receive signals from the same or different modalities but do *not* map sensory or motor activity in a way that preserves feature-based, topographic and topological relations of the external environment as they appear in psychological experience. Convergence zones do not embody a refined representation, in the sense that would be assumed in an information-processing model, although they do route information in the sense of information theory. They know "about" neural activity in the feeding cortices and can promote further cortical activity by feedback/retroactivation. In themselves, however, they are uninformed as to the content of the representations they assist in attempting to reconstruct. The role of convergence zones is to enact formulas for the reconstitution of fragment-based momentary representations of entities or events in sensory and motor cortices—the experiences we remember.³

2. *Operating principles*

Convergence zones signal the related binding of the similarity, spatial placement, temporal sequence, or temporal coincidence of feature-based fragments highlighted in the perceiver's experience. Convergence zones prompt sensory and motor co-activation by means of back-projections into cortices located upstream. In the extreme view (a mere caricature), all that would be required of a convergence zone would be to function as a pivot, that is, to cause retroactivation in sites that it fed back to, after a threshold defined by concurrent inputs had been reached. The general operating principle would be stated as: (a) reactivate itself when fired upon; (b) reactivation promotes firing toward any site to which there are back-projections, recip-

³The notion of separating storage of fragments of experience, from storage of a catalogue for their reconstitution, was inspired by our study of patient Boswell, along with the notion that a unidirectional caudal-rostral processing cascade was less likely than a multidirectional, recursive organization. The idea of convergence zones came from reflection on patterns of cortico-limbic projections, especially the multiplicity of parallel and converging channels, and the progressive size reduction of the neural convergence sites along a caudal-rostral axis. The pattern of disruption of cortico-limbic and cortico-cortical feed-forward and feedback projections in patients with Alzheimer's disease (see Van Hoesen & Damasio, 1987, for a review) provided the blueprint for the construct.

roccating feed-forward inputs that generated the synaptic pattern that defines the zone. But because of superimposition and overlapping of convergence zones within the same neuron population, and of the ensuing high number of synaptic interactions, the range of back-firing of each convergence zone is modulated rather than rigid. It depends on the momentary number and nature of cortical feed-forward inputs (relative to the total number of possible feedback outputs that the zone can have), and on the momentary inputs from other areas of cortex and from limbic system, thalamus, basal forebrain, and so forth.

As a consequence, convergence zones can produce different ranges of retroactivation in the cortex, depending on the concurrent balances of inputs they receive. Also, convergence zones can blend responses, that is, produce retroactivation of fragments that did not originally belong to the same experiential set, because of underspecification of cortical feed-forward inputs, or higher-order cortical feedbacks, or subcortical feedbacks. When pathological combinations of input are reached, the zone malfunctions, for example, it may generate "fantastic" or "psychotic" responses, or not operate at all.

It is important to note that the lines activated by feedback from convergence zones are not rigid. They should be seen as facilitated paths that may or may not be travelled depending on the ensemble pattern of synaptic interactions within a population.

3. Types of convergence zones

I envisage permanent convergence zones in the cortex and temporary convergent zones in limbic structures and basal ganglia/cerebellum, based on current findings regarding the profile of retrograde amnesia following hippocampal damage. The domain of the convergence zone is determined by its immediate and remote feed-forward inputs which are co-extensive with its back-projection targets.

I propose two types of convergence zones. In Type I, the zone fires back simultaneously and produces concomitant activations. Type I zones inscribe temporal coincidences and aim at replicating them. Type II convergence zones fire back in sequence, producing closely ordered activations in the target cortices. Such zones have inscribed temporal sequences and aim at replicating them. The time scale for firing from Types I and II convergence zones would be different.

Type I convergence zones are located in sensory association cortices of low and high order, and are assisted in learning by the hippocampal system. Type II convergence zones are the hallmark of motor-related cortices, and are assisted in learning by basal ganglia and cerebellum.

In the normal condition, the two types of convergence zone interlock at

multiple levels so that learning relative to an entity or event recruits both types of convergence zones. Likewise normal recall and recognition involve operations in both types of convergence zone, even when the triggering stimulus only activates one type of convergence zone at the outset of the process.

4. The development of convergence zones

The placement of convergence zones is partly the result of the genetically expressed neuroanatomical design and partly the result of the sculpting process introduced by learning. Convergence zones develop in association cortices that: (a) receive projections in a convergent manner from a wider array of cortices located upstream; (b) can reciprocate projections to the feeding cortices; (c) can project downstream to other cortices and subcortical structures; and (d) can receive a wide array of projections from several subcortical and motor structures.

It is the genetic pattern of neuroanatomical connections that first constrains the potential domain of convergence zones. For example, a convergence zone in early visual association cortices cannot possibly bind anything but visually related activity at the level of component features, whereas a convergence zone in anterior temporal cortices can be told about activity related to numerous simultaneous events and bind their coincidence. But the ultimate anatomical location and functional destiny of convergence zones is determined by learning, as neuron ensembles become differentially dedicated to certain types of occurrence in feeding cortices.

Convergence zones are created during learning as a result of concurrent activations in neuron ensembles within association cortices of different order, hippocampus, amygdala, basal ganglia, and cerebellum. The concurrent activations come from convergent feed-forward signals generated by neural activity in: (a) sensory and motor cortices (as caused by perception or recall of external or internal entities); (b) feedback projections from other convergence zones in association cortices; (c) direct and indirect feedback projections from convergence zones in limbic cortices and from limbic related nuclei; (d) direct and indirect feedback projections from basal ganglia, non-motor thalamus, and cerebellum; and (e) local microcircuitry interactions.

As noted above, convergence zones have thresholds and levels of response. The activation of a convergence zone depends on its internal constitution, the size, locus, number, and location of sensory and motor representation sites that it subtends. It also depends on the momentary concurrent combination of potential trigger weights, from neural activity related to externally generated representations, internally recalled representations, and back-projection from all the neuronal sites listed previously.

5. *Superposition of signals*

Convergence zones contain overlapping binding codes for many entities and events. Such rich binding is the source of the widening retroactivation that permits recognition and thought processes, and yet its wealth, if unchecked, would eventually result in co-activations bearing only minimal relationships to previous specific experiences and on inability to reconstitute unique events. Ultimately, fantastic and cognitively catastrophic combinations would occur, as they do in fact occur in a variety of neuropsychological disorders caused by the neuropathological processes at several levels of the system. In the normal brain, the constraints that impose specificity of co-evocations depend on concurrent inputs from the following systems: (a) other convergence zones, at multiple neural levels, whose subtended retroactivation provides neural context and thereby helps constrain co-activation; and (b) non-specific limbic nuclei (basal forebrain and brain stem) activated by antero-temporal limbic units (amygdala, hippocampus).

6. *Attention*

In a system that produces multiple-site activations incessantly, it is necessary to enhance pertinently linked sites in order to permit binding by salient coincidences. I use the term attention to designate the “spotlighting” process that generates simultaneous and multiple-site salience and thus permits the emergence of evocations. Consciousness occurs when multiple sites of activation are simultaneously enhanced in keeping or not with real past experiences. (Some psychotic and demential states are possibly examples of simultaneous enhancement of activations whose combination does not conform to reality; in non-pathological states the same applies to day-dreams). As defined here, attention depends on numerous factors and mechanisms. First, there is a code for enhancement of activations that is part of the record of the activation pattern it enhances. Type II convergence zones are especially suited to this role. Secondly, the state of the perceiver and the context of the process play important roles in determining the level of activations. The reticular activating system, the reticular complex of the thalamus, and the limbic system mediate such roles under partial control of the cerebral cortex.

The evocations that constitute experienced recall occur in specified sensory and motor cortices, albeit in parcellated fashion. Experienced recall thus occurs where physical structures of external entities or body states were mapped in feature fragment manner, notwithstanding the fact that a complex neural machinery made up of numerous other areas of cortex and subcortical nuclei cooperates to reconstruct the co-activation patterns and enhance them.

7. The placement of convergence zones

Convergence zones have different placements within association cortices and other gray matter regions, and varied activation thresholds. There are numerous levels of convergence zone depending on knowledge domain and contextual complexity (taxonomic level). The functional regionalization of a domain corresponds to the neural inscription of separate sensory and motor activities related to features and dimensions of different exemplars. The inscriptions are naturally superimposed to the extent that the respective features and dimensions overlap, or coincide in time. The inscriptions are naturally contiguous when the respective features or acts they represent occurred in temporal sequence. As superimpositions accrue, categories emerge from the blends and mergings of separate exemplars. It is important to note that for each separate exemplar to be recalled as an individual entity, it is necessary to add contextual complexity to its representation. This is accomplished by connecting its inscription to the inscription of other entities and events so that an entirely unique set can be defined. When additional inscriptions are not linked to create unique or nearly unique sets, the superimposition of exemplars remains categorical or generic, and recall can reconstitute any one previously learned exemplar or else a blend of exemplars. The creation of records of contextual complexity, which code for the temporal entities and events, is thus critical for recall or recognition at unique (episodic) level.

It is important to note that in this perspective the building of categories occurs while inscribing episodes. The system operates so that it always attempts to inscribe as much as possible of the entire context. Even if the system fails to inscribe the whole episode—or if it does inscribe it, but recall cannot fully reconstitute it—the operation preserves enough of the core inscription of an entity (or event) for categorization to develop from this and other related inscriptions. The inscription of categories precedes episode inscription; that is, it is neuroanatomically and neurophysiologically more caudal. This disposition explains the impairment of episodic memory and preservation of generic memory following damage to anterior temporal cortices.

Knowledge of objects, faces, numbers, among many others, created by perceptuomotor interactions, is anatomically and functionally regionalized in a manner different from classic localizationism of function, but that does admit a notable degree of anatomical specialization. This form of specialization does not follow traditional anatomical boundaries such as are known for sensory modalities, or cytoarchitectonic brain areas. Nor does it conform to the functional centers of traditional neurology. The fragment representations that comprehensively describe an entity are dispersed by multiple functional regions which are, in turn, located in different cytoarchitectonic areas. The

many convergence zones necessary to bind the fragments relationally are located in yet other neural sites. The region thus formed obeys anatomical criteria dictated by the nature of the entity represented, and by the interaction between perceiver and entity, and is secondarily constrained by the potential offerings of the anatomy. The comprehensive representation of a specific entity or category is distributed not only within a population of neurons but is also distributed in diverse types of neural structure, cortically and subcortically. In this proposal, the term localization can only refer to an imaginary space defined by neural sites likely to contain convergent zones necessary for the retroactivation of a given set of entities or events. The borders of such a space are not only fuzzy but changeable, in the sense that for different instantiations of retroactivation of a given entity the set of necessary convergence zones varies considerably.

Applications of the framework

In the following two sections I discuss briefly the application of this proposal to learning and memory and language.

Learning and memory

1. The relative segregation of memory domains

The fact that different neural regions support memory for different domains is the reason why striking performance dissociations can occur in human amnesia. For instance, after lesions in the hippocampal system, patients retain previously learned perceptuomotor skills (so-called procedural knowledge) or even learn new ones, while memory and learning for new faces or objects is no longer possible (Cohen & Squire, 1980; Damasio et al., 1985a,b, 1987; Eslinger & Damasio, 1986; Milner, Corkin, & Teuber, 1968). This dissociation occurs because the representations of motor entities rely on structures that remain intact in those patients: somatosensory and motor cortices, neo-striatum and cerebellum. As noted above, the functional essence behind the system formed by those structures is the recording and re-enactment of temporal sequences and relies on Type II convergence zones.

Participation of the hippocampal system is not at all necessary for the acquisition and maintenance of procedural memories, provided they are used only at a covert level, and the subject is not required to recollect the factual information related to the acquisition of the skill or to the circumstances in

which the skill has been previously exercised. Conscious recall of the source of knowledge requires patency of at least one hippocampal region.

By contrast, the weight of recording factual knowledge, in spite of its diverse base on sensory and motor activities, relies most importantly on sensory cortices and necessitates hippocampal activity. The functional essence in this system is the recording of neural activity related to physical structure (of features, entities, and events), spatial contiguity (of features and entities), and temporal coincidence (of entities and events). Type I convergence zones in the hippocampal-bound association cortices are required. Perhaps the most dramatic lesion-related dissociation within factual knowledge is the one that compromises memory for complex social events but spares general knowledge of entities and events outside of a social context (Damasio & Tranel, 1988; Eslinger & Damasio, 1985). Other striking dissociations abound, however, for different categories of objects, for verbal and non-verbal knowledge, and for different types of verbal knowledge (Damasio et al., 1989b).

2. Different levels of memory processing

In essence, the distinction between generic and episodic memories is a distinction of processing levels during recall or recognition. We can recall at generic levels, with little contextual complexity attached to an entity, no definition of uniqueness, and no connection to our autobiography. Or we can recall at progressively richer episodic levels, with the evocation of greater contextual complexity and the experience of autobiographic events in which entities play more specific roles. I believe the brain normally attempts to capture the maximal complexity of every event, although the stability of the recording of such complexity varies with the value of the event and with the anticipated need to recall it.

3. The mapping of uniqueness and of entity-centered knowledge

The critical distinction between generic and episodic knowledge, from the standpoint of learning, resides with the ability to record temporal coincidence (co-occurrence) of entities within a wide and complex context. It is a matter of magnitude that distinguishes generic from episodic levels of processing, somewhat artificially, along a continuum.

When a perceiver interacts with a novel entity, learning consists of recording any additional patterns of physical structure, somatic state, or relational binding that transpired during the interaction but were *not previously recorded*. The same applies to learning of new events.

In virtually all instances of learning beyond the early acquisition periods of infancy and childhood, any new pattern of activity related to perception

of new entities and events also evokes multiple and previously stored patterns that are thus co-experienced with the novel stimuli. Learning does not entail the recording of all the information contained in a new event. Rather, it calls for the co-evocation of many physical structures and relations previously recorded for related events, the recording of any novel features that had not been recorded before, and the linking of novel records with the pre-existing records so that a new specific set is defined and the code for its potential reconstitution committed to a convergence zone.

There is a large sharing of memory records such that the same neural patterns can be applied to many entities and events by superimposition and overlap whenever and wherever their physical structure or relational bindings are shared. The inscription of a specific entity or event can be made unique only by means of connecting a particular component to others. Such an organization is extremely economical and promotes a large memory capacity. However, it is also prone to ambiguity and an easily disordered operation if one of its many supporting devices malfunctions. Confusional states and some amnesic syndromes caused by subcortical lesions are an expression of such malfunctions. At a milder level, fatigue, sleep deprivation, or distraction can cause the same.

4. Neural substrates for learning and memory at systems level

The critical neural substrate for learning and memory comprises two major subsystems: one that interconnects sensory cortices assigned to mapping physical structure and temporal coincidence with the hippocampus; and a second that interconnects sensory and motor cortices assigned to mapping temporal sequence with the basal ganglia/cerebellum and the dorsolateral frontal cortices. Normal operation of these subsystems is cooperative rather than independent.

The neuroanatomical design of the entorhinal cortex and of the sequence of cellular regions in the hippocampus to which it projects deserves special mention. This subsystem provides a set of auto-interacting convergence zones of great complexity. It is the only brain region in which signals originally triggered by neural activity in all sensory cortices and in centers for autonomic control can actually co-occur over the same neuron ensembles. As such, this is the appropriate substrate for a detector of temporal coincidences, the function that I have previously proposed for this system and that I believe to be lost in amnesia following hippocampal damage (see Damasio et al., 1985a). Such a function is compatible, in essence, with the type of physiological basis for learning proposed by Hebb, a presynaptic/postsynaptic coincidence mechanism. It is also compatible with a variety of recent cellular and molecular evidence regarding the phenomenon of long-term potentiation (LTP) and

the role of NMDA-gated calcium channels as detectors of coincidence (see Gustafsson & Wigstrom, 1988, for a review).

Once detection of co-occurrence takes place, the region acts via its powerful feedback system into cortical and subcortical neural stations, to assist in the creation or modification of convergence zones located in the cortices that originally projected into the entorhinal cortex. It is also apparent that such a structure, especially the autocorrelation matrix of CA₃, could store within itself binding codes of the kind I envisage for convergence zones, capable of content-addressed completion. It appears unlikely, however, that the hippocampal complex remains as a storage site for long periods, not only because of what that would mean in terms of capacity limits and risk of malfunction, but also because bilateral damage confined to the entorhinal cortex/hippocampus appears to cause only limited impairments of retrograde memory (Corkin, 1984), and the same appears to be true of bilateral damage to the hippocampus alone (Zola-Morgan, Squire, & Amaral, 1986). The definitive account on this issue is not available yet. In humans, the left and right hippocampi appear to be dedicated to different operations and may also operate differently in terms of their long-term role in retrieval.

5. Consciousness and self-consciousness

As previously noted, consciousness emerges when retroactivations attain a level of activity that confers salience. Coincident salient sites of activity define a set that separates itself from background activity and emerges, in psychological terms, as a conscious content on evocation as opposed to non-salient retroactivations that remain covert.

Conscious contents are all contents about which one can give testimony, in verbal narrative form, but I wish to distinguish them from the subset of conscious content we call self-conscious contents. The difference resides with the notion of self and autobiography. In my view, self-consciousness only emerges when conscious contents relative to an ongoing stimulus are experienced in the context of pertinent autobiographical data. The distinction is not specious. Patient Boswell is conscious of his environment and properly recognizes the stimuli around him but not in relation to his autobiography. Whether the stimulus is something that he ought to have recognized as unique, or something truly new to him, his ability to put it in the perspective of his life experience is restricted. His self-consciousness is thus limited and unlike that of perceivers in whom evocations generated by novel percepts are co-attended simultaneously with autobiographical evocations.

Language

The representations related to language, that is, the representations of lexical entries and grammatical operations, including syntactic rules or principles, phonology, morphology, and semantics which constitute the internalized or mental grammar, are perceived, acquired, and co-activated according to the principles articulated for non-verbal entities. As noted above, the framework does not address the issue of innate versus acquired aspects of language, although from a perspective of biological evolution as well as from the investigation of universal properties of the world's diverse languages it is likely that the substrates for combinatorial semantics and syntactical principles are partly innate.

The lexicon and language-specific aspects of the grammar, as cultural artifacts, are a subset of reality characterized by certain physical structures (the physical phonetic articulatory gestures and resultant acoustic correlates of linguistic units and structures, that is, phones, phonemes, morphemes, words, phrases, sentences, etc.) and logical relationships (grammatical functions) at multiple levels. Those external physical structures and relations constitute a corpus of signals capable of symbolizing, in sentential terms, most non-language aspects of reality at any level. By means of both feature-based physical fragment representations and binding convergence zones, the brain stores the potential for reconstituting any lexical entry or relational arrangement that it has learned, as well as the implicit rules by which novel utterances are produced and comprehended. This would not deny the possibility that highly frequent lexical entries would be recorded at large-scale fragment level, for instance, the level of an entire word stem, a condition that would be highly adaptive.

The brain not only inscribes language constituents but also provides direct and dynamic neural links between verbal representations and the representation of non-language entities or events that are signified by language. In other words, the brain embodies (materializes) in neural hardware the combined biological and cultural bond that culture has assigned between a language representation (a signifier) and a segment of non-verbal reality (a signified), to borrow Saussure's suggestive terminology. It is that neural bond that permits the two-way, uninhabitable translation process that can automatically convert non-verbal co-activation into a verbal narrative (and vice versa), at every level of neural representation and operation.

Testing the framework

There are fundamentally four approaches to test the validity of the hypotheses expressed here. One relies on the lesion method, the approach on which

most of these ideas are based. Small focal and stable lesions in humans with neurological disease can be used to probe neuropsychological predictions based on the hypotheses expressed here. Another approach involves the use of positron emission tomography in both normals and patients with focal brain damage, to explore temporal correlations among different cortical regions activated by controlled stimuli. Another approach would involve computational modeling and testing of the concept of convergence zone. Finally, it will be possible in experiments using multiple recording from different cortical sites to test the notion of time-locked activations. For instance, in an experiment where one would record simultaneously from multiple cortical sites encompassing two sensory modalities, the following should be observed:

- (1) After a delay compatible with feedback firing, electrical stimulation of convergence zones would produce synchronous activity in separate cortical sites presumed to contain feature-fragments related to the convergence zone.

The regions chosen for stimulation would be guided by knowledge of neurons in association cortex that respond constantly to specific stimuli, for example, faces. Likewise, the choice of areas to guide the search for time-locked activity in early cortices would come from knowledge of areas known to be activated by the perception of a specific stimulus, for example, a face.

- (2) The lack of finding of time-locked activity across a vast array of cortical regions theoretically presumed to be necessary for the reconstitution of a perceptual set would constitute evidence against the notion of convergence zones proposed here.

Situating the proposal

I see the following features of the theory as distinctive:

- (1) The notion that there is a major distinction between records of physical structure fragments, and records of combinatorial arrangements among those records.
- (2) The notion that the experience of entities or events in recall always depends on the time-locked retroactivation of fragmentary records contained in multiple sensory and motor regions and thus on momentary attempted reconstitutions of the once perceived components of reality.
- (3) The notion that while evocations only exist momentarily, they are the only directly inspectable aspect of brain activity. Their fleeting existence makes them no less real. Furthermore, although their existence depends

on a complex machinery distributed by multiple brain sites and levels, the proposal specifies that the attempted reconstitutions occur in an anatomically restricted sector of the cerebrum.

- (4) The notion that certain aspects of the interaction between perceiver and reality generate domains of knowledge, which become regionalized according to neural constraints rather than conceptual-lexical labels.
- (5) The notion that the anatomical placement and connectional definition of a convergence zone, that is, the specification of its inputs and outputs at the point in the system that is located, also defines the knowledge domain the convergence zone embodies.
- (6) The role attributed to feedback projections, especially cortico-cortical, in the mechanisms of reconstitution of experiences. Feedback is distinguished from re-entry as used in the automata of Edelman and Reeke (1982). Feedback and feed-forward carry signals about activity in interconnected units but they do not transport a movable representation being entered or re-entered. Feed-forward signals mark the presence of activity upstream in the network, and indicate the whereabouts of records of activity. Feedback reactivates such upstream records. The convergence zones record those relationships and operate to route activity. No representations of reality as we experience it are ever transferred in the system; that is, no concrete contents and no psychological information move about in the system.
- (7) The value accorded to representations of internal somatic states in all their aspects and levels. Somatic states are generally relegated to a subsidiary position, a matter of non-specific influence on the general workings of a network concerned with representations of external reality. In this proposal somatic states are memorized in feature-based fragment records (linked by binding convergence zones), just as external stimuli are. The source for this notion was our studies of humans with focal lesions, especially those with conditions such as anosognosia and acquired disorders of conduct (Damasio & Anderson, 1989).

It is perhaps useful to compare this proposal to other recent proposals that deal with cognitive processes and the organization of their putative neural substrates. In order to do this we will choose two reference points: the classical model of cognitive architecture, as presented, for instance, by Fodor and Pylyshyn (1988), and a range of models known under the designations "parallel distributed processing" or "connectionism" (see Rumelhart & McClelland, 1986).

We believe that the structures and operations described in this theory occupy an intermediate position and are compatible with the proposals in

these reference points. The neural organization we propose is at the level of systems formed by macroscopic functional regions. It embodies and can implement some predicates of a classical cognitive architecture. On the other hand, it is conceivable that connectionist nets and algorithms may realize some of the microscopic levels underlying the organization proposed here. By the same token our theory is also compatible with neuronal group selection theory (Edelman & Finkel, 1984). Although the specification of neuron units in those theories is designed in “brain-style”, the overall networks are not yet “brain-like”. The principles of structure and operation of the machines so designed are not aimed at the superstructure organization necessary for cognitive processes such as thought, language, or consciousness; that is, to our knowledge they do not yet compel separate units to hook themselves up in a particular way capable of making a system thoughtful and self-conscious. By contrast, this theory seeks to propose precisely some of those higher organization principles. Cognitive architecture proposals refer to psychological phenomena that our framework aims at capturing. Connectionist models refer to microstructure and function situated below the levels at which our description concentrates, but that might conceivably carry on some of the necessary implementations, at least in certain sectors of the neural structure.

References

- Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience*, 8, 407–430.
- Barlow, H.B. (1981). Critical limiting factors in the design of the eye and visual cortex. (The Ferrier Lecture, 1980). *Proceedings of the Royal Society London, (Biology)*, 212, 1–34.
- Bauer, R.M. (1984). Autonomic recognition of names and faces in prosopagnosia: A neurophysiological application of the Guilty Knowledge Test. *Neuropsychologia*, 22, 457–469.
- Bruce, C.J., Desimone, R., & Gross, C.G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 46, 369–384.
- Chavis, D.A., & Pandya, D.N. (1976). Further observations on corticofrontal connections in the rhesus monkey. *Brain Research*, 117, 369–386.
- Cohen, N.J., & Squire, L.R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210, 207–210.
- Corkin, S. (1984). Lasting consequences of bilateral medial temporal lobectomy: Clinical course and experimental findings in HM. *Seminars in Neurology*, 4, 249–259.
- Crick, F. (1984). Function of the thalamic reticular complex: The searchlight hypothesis. *Proceedings of the National Academy of Science USA*, 81, 4586–4590.
- Damasio, A.R. (1979). The frontal lobes. In K. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology*. New York: Oxford University Press.
- Damasio, A.R. (1985a). Disorders of complex visual processing. In M.M. Mesulam (Ed.), *Principles of behavioral neurology*. Philadelphia: Davis.

- Damasio, A.R. (1985b). The frontal lobes. In K. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (2nd edition). New York: Oxford University Press.
- Damasio, A.R. (1989). Category-related recognition defects as a clue to the neural substrates of knowledge. *Trends in Neurosciences*, in press.
- Damasio, A.R., & Anderson, S. (1989). Anosognosia as a domain-specific memory defect. *Journal of Clinical and Experimental Neuropsychology*, *11*, 17.
- Damasio, A.R., & Damasio, H. (1983). The anatomic basis of pure alexia. *Neurology*, *33*, 1573–1583.
- Damasio, A., Damasio, H., & Tranel, D. (1989a). Impairments of visual recognition as clues to the processes of memory. In G.M. Edelman, W.E. Gall, & W.M. Cowan (Eds.), *Signal and sense: Local and global order in perceptual maps*. In press.
- Damasio, A.R., Damasio, H., & Tranel, D. (1989b). New evidence in amnesic patient Boswell: Implications for the understanding of memory. *Journal of Clinical and Experimental Neuropsychology*, *11*, 61.
- Damasio, A., Damasio, H., Tranel, D., Welsh, K., & Brandt, J. (1987). Additional neural and cognitive evidence in patient DRB. *Society for Neuroscience*, *13*, 1452.
- Damasio, A., Damasio, H., & Van Hoesen, G.W. (1982). Prosopagnosia: Anatomic basis and behavioral mechanisms. *Neurology*, *32*, 331–341.
- Damasio, A., Eslinger, P., Damasio, H., Van Hoesen, G.W., & Cornell, S. (1985a). Multimodal amnesic syndrome following bilateral temporal and basal forebrain damage. *Archives of Neurology*, *42*, 252–259.
- Damasio, A., Graff-Radford, N., Eslinger, P., Damasio, H. & Kassell, N. (1985b). Amnesia following basal forebrain lesions. *Archives of Neurology*, *42*, 263–271.
- Damasio, A.R., & Tranel, D. (1988). Domain-specific amnesia for social knowledge. *Society for Neuroscience*, *14*, 1289.
- Desimone, R., Albright, T.D., Gross, C.G., & Bruce, C. (1984). Stimulus-selective responses of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*, 2051–2062.
- Desimone, R., Schein, S.J., Moran, J., & Ungerleider, L.G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Research*, *25*, 441–452.
- Desimone, R., & Ungerleider, L. (1989). Neural mechanisms of visual processing in monkeys. In A. Damasio (Ed.), *Handbook of neuropsychology: Disorders of visual processing* (Vol. II, pp. 267–299). Amsterdam: Elsevier.
- Edelman, G.M., & Finkel, L.H. (1984). Neuronal group selection in the cerebral cortex. In G.M. Edelman, W.E. Gall, & W.M. Cowan (Eds.), *Dynamic aspects of neocortical function*. New York: Wiley.
- Edelman, G.M., & Reeke, G.N. (1982). Selective networks capable of representative transformations, limited generalizations, and associative memory. *Proceedings of the National Academy of Science USA*, *79*, 2091–2095.
- Eslinger, P.J., & Damasio, A.R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation. *Neurology*, *35*, 1731–1741.
- Eslinger, P.J., & Damasio, A.R. (1986). Preserved motor learning in Alzheimer's disease. *Journal of Neuroscience*, *6*, 3006–3009.
- Farah, M. (1989). The neuropsychology of mental imagery. In A. Damasio (Ed.), *Handbook of neuropsychology: Disorders of visual processing*. (Vol. II, pp. 395–413). Amsterdam: Elsevier.
- Feinberg, T., Rothi, L., & Heilman, K. (1986). Multimodal agnosia after unilateral left hemisphere lesion. *Neurology*, *36*, 864–867.
- Fodor, J.A., & Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, *88*, 237–294.
- Geschwind, N., & Fusillo, M. (1966). Color-naming defects in association with alexia. *Archives of Neurology*, *15*, 137–146.
- Gilbert, C.D. (1983). Microcircuitry of the visual cortex. *Annual Review of Neuroscience*, *6*, 217–247.
- Goldman-Rakic, P.S. (1984). The frontal lobes: Uncharted provinces of the brain. *Trends in Neurosciences*, *7*, 425–429.

- Goldman-Rakic, P.S. (1988). Topography of cognition: Parallel distributed networks in primate association cortex. In: *Annual Review of Neuroscience*, 2, 137–156.
- Gustafsson, B., & Wigstrom, H. (1988). Physiological mechanisms underlying long-term potentiation. *Trends in Neurosciences*, 11, 156–162.
- Horestein, S., Chamberlin, W., & Conomy, J. (1967). Infarction of the fusiform and calcarine regions: Agitated delirium and hemianopia. *Transactions of the American Neurological Association*, 92, 85–89.
- Hubel, D.H., Livingstone, M.S. (1987). Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience*, 7, 3378–3415.
- Hubel, D.H., & Wiesel, T.N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society London Series B*, 198, 1–59.
- Jones, E.G., & Powell, T.P.S. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*, 93, 793–820.
- Julesz, B. (1971). *Foundation of cyclopean perception*. Chicago: University of Chicago Press.
- Julesz, B. (1981). Textons, the elements of texture perception and their interaction. *Nature*, 290, 91–97.
- Kosel, K.C., Van Hoesen, G.W., & Rosene, D.L. (1982) Nonhippocampal cortical projections from the entorhinal cortex in the rat and rhesus monkey. *Brain Research*, 244, 202–214.
- Kosslyn S.M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- Lettvin, J.Y., Maturana, H.R., McCulloch, W.S., & Pitts, W.H. (1959). What the frog's eye tells the frog's brain. *Proceedings of the IRE*, 47, 1940–1949.
- Lewis, D.A., Campbell, M.J., Foote, S.L., & Morrison, J.H. (1986). The monoaminergic innervation of primate neocortex. *Human Neurobiology*, 5, 181–188.
- Livingstone, M.S., & Hubel, D.H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, 4, 309–356.
- Livingstone, M.S., & Hubel, D.H. (1987a). Connections between layer 4B of area 17 and thick cytochrome oxidase stripes of area 18 in the squirrel monkey. *Journal of Neuroscience*, 7, 3371–3377.
- Livingstone, M.S., & Hubel, D.H. (1987b). Psychological evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, 7, 3416–3468.
- Livingstone, M.S., & Hubel, D.H. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, 740–749.
- Lund, J.S., Hendrickson, A.E., Ogren, M.P., & Tobin, E.A. (1981). Anatomical organization of primate visual cortex area VII. *Journal of Comparative Neurology*, 202, 19–45.
- Luria, A.R. (1966). *Higher cortical functions in man*. New York: Basic Books.
- McCarthy, R.A., & Warrington, E.K. (1988). Evidence for modality-specific meaning systems in the brain. *Nature*, 334, 428–430.
- Mesulam, M.M., Mufson, E.J., Levey, A.I., & Wainer, B.H. (1983). Cholinergic innervation of the cortex by basal forebrain: Cytochemistry and cortical connections of the septal area, diagonal band nuclei, nucleus basalis (substantia innominata) and hypothalamus in the rhesus monkey. *Journal of Comparative Neurology*, 214, 170–197.
- Milner, B., Corkin, S., & Teuber, H.L. (1968). Further analyses of the hippocampal amnesic syndrome: 14 year follow-up study of H.M. *Neuropsychologia*, 6, 215–234.
- Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J.L. McGaugh, & N.M. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 65–77). New York: Guilford Press.
- Mountcastle, V.B., Lynch, J.C., & Georgopoulos, A. (1975). Posterior parietal association cortex of the monkey: Command functions for operations within extra-personal space. *Journal of Neurophysiology*, 38, 871–908.
- Nauta, W.J.H. (1971). The problem of the frontal lobe: A reinterpretation. *Journal of Psychiatric Research*, 8, 167–187.

- Newcombe, F., & Ratcliff G. (1974). Agnosia: A disorder of object recognition. In F. Michel & B. Schott (Eds.), *Les syndromes de disconnexion calleuse chez l'homme*. Lyon: Colloque international de Lyon.
- Paillard, J. (1971). Les déterminants moteurs de l'organisation de l'espace. *Cahiers de Psychologie*, 14, 261–316.
- Pandya, D.N., & Kuypers, H.G.J.M. (1969). Cortico-cortical connections in the rhesus monkey. *Brain Research*, 13, 13–36.
- Pandya, D.N., & Yeterian, E.H. (1985). Architecture and connections of cortical association areas. In A. Peters & E.G. Jones (Eds.), *Cerebral cortex*, (Vol. 4). New York: Plenum Press.
- Penfield, W., & Jasper, W. (1954). *Epilepsy and the functional anatomy of the human brain*. Boston: Little, Brown.
- Perrett, D.I., Mistlin, A.J., & Chitty, A.J. (1987). Visual neurons responsive to faces. *Trends in Neurosciences*, 10, 358–364.
- Perrett, D.I., Rolls, E.T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329–342.
- Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Rockland, K.S., & Pandya, D.N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Research*, 179, 3–20.
- Rockland, K.S., & Pandya, D.N. (1981). Cortical connections of the occipital lobe in rhesus monkey: Interconnections between areas 17, 18, 19 and the superior temporal gyrus. *Brain Research*, 212, 249–270.
- Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Rosene, D.L., & Van Hoesen, G.W. (1977). Hippocampal efferents reach widespread areas of the cerebral cortex in the monkey. *Science*, 198, 315–317.
- Rumelhart, D.E., & McClelland, J.L. (1986). *Parallel distributed processing (Vol. 1)*. Cambridge, MA: MIT Press.
- Sejnowski, T.J. (1986). Open questions about computation in cerebral cortex. In J.L. McClelland & D.R. Rumelhart (Eds.), *Parallel distributed processing* (pp. 372–389). Cambridge, MA: MIT Press.
- Seltzer, B., & Pandya, D.N. (1976). Some cortical projections to the parahippocampal area in the rhesus monkey. *Experimental Neurology*, 50, 146–160.
- Seltzer, B., & Pandya, D.N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Research*, 149, 1–24.
- Squire, L.R. (1987). *Memory and brain*. New York: Oxford University Press.
- Swanson, L.W., & Kohler, C. (1986). Anatomical evidence for direct projections from the entorhinal area to the cortical mantle in the rat. *Journal of Neuroscience*, 6, 3010–3023.
- Tranel, D., & Damasio, A. (1985). Knowledge without awareness: An autonomic index of facial recognition by prosopagnosics. *Science*, 228, 1453–1454.
- Tranel, D., & Damasio, A. (1987). Autonomic (covert) discrimination of familiar stimuli in patients with visual agnosia. *Neurology*, 37, 129; *Society for Neuroscience*, 13, 1453.
- Tranel, D., & Damasio, A. (1988). Nonconscious face recognition in patients with face agnosia. *Behavioral Brain Research*, 30, 235–249.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press.
- Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, R.J.W. Mansfield, & M.A. Goodale (Eds.), *The analysis of visual behavior*. Cambridge, MA: MIT Press.
- Van Essen, D.C. (1985). In A. Peters & E.G. Jones (Eds.), *Functional organization of primate visual cortex*. New York: Plenum Publishing.

- Van Essen, D.C., & Maunsell, J.H.R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, 6, 370–375.
- Van Hoesen, G.W., & Damasio, A.R. (1987). Neural correlates of the cognitive impairment in Alzheimer's disease. In F. Plum (Ed.), *Higher functions of the nervous system: The handbook of physiology* (pp. 871–898).
- Van Hoesen, G.W., & Pandya, D.N. (1975a). Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices in the rhesus monkey. I. Temporal lobe afferents. *Brain Research*, 95, 1–24.
- Van Hoesen, G.W., & Pandya, D.N. (1975b). Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices in the rhesus monkey. III. Entorhinal cortex efferents. *Brain Research*, 95, 39–59.
- Van Hoesen, G.W., Pandya, D.N., & Butters, N. (1975). Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices in the rhesus monkey. II. Frontal lobe afferents. *Brain Research*, 95, 25–38.
- Van Hosen, G.W. (1982). The primate parahippocampal gyrus: New insights regarding its cortical connections. *Trends in Neurosciences*, 5, 345–350.
- Zeki, S.M. (1987). Personal communication.
- Zola-Morgan, S., Squire, L.R., & Amaral, D. (1986). Human amnesia and the medial temporal region: Enduring memory impairment following a bilateral lesion limited to the CA1 field of the hippocampus. *Journal of Neuroscience*, 6, 2950–2967.

Résumé

Cet article esquisse un cadre théorique pour la compréhension des bases neurales de la mémoire. Il propose l'hypothèse que le rappel et la reconnaissance d'entités et d'événements dépendent de l'activation de nombreux ensembles de neurones dans de multiples régions des cortex sensoriels et moteurs où les représentations des fragments de formes sont représentés par des configurations d'activité impliquées à l'origine par les interactions perceptuo-motrices. Le processus d'activation est dirigé à partir de multiples zones de convergence situées dans les cortex d'association et dans certains noyaux gris sous-corticaux. Les zones de convergence enregistrent de façon amodale l'arrangement combinatoire des différents fragments de formes tels qu'il se présente dans les aires corticales précoces au cours de l'entité ou de l'événement. Les zones de convergence sont reliées avec les ensembles neuronaux primaires par des projections réciproques qui forment des chemins facilités plutôt que des liens rigides. Le fonctionnement des zones de convergence est modulé de façon dynamique par les entrées concurrentes provenant d'autres aires et noyaux sous-corticaux. Ce modèle refuse l'existence d'un site anatomique unique pour l'intégration sensori-motrice et d'une mémoire unique gardant le sens d'entités ou d'événements. Le sens résulte de la rétro-activation distribuée et synchrone de fragments. Seuls ces derniers atteignent le seuil du conscient.