

The Brain Binds Entities and Events by Multiregional Activation from Convergence Zones

Antonio R. Damasio

Department of Neurology,

Division of Behavioral Neurology and Cognitive Neuroscience,

University of Iowa College of Medicine, Iowa City, IA, USA

The experience of reality, in both perception and recall, is spatially and temporally coherent and “in-register.” Features are bound in entities, and entities are bound in events. The properties of these entities and events, however, are represented in many different regions of the brain that are widely separated. The degree of neural parcellation is even greater when we consider that the perception of most entities and events also requires a motor interaction on the part of the perceiver (such as eye movements and hand movements) and often includes a recordable modification of the perceiver’s somatic state. The question of how the brain achieves integration starting with the bits and pieces it has to work with, is the binding problem. Here we propose a new solution for this problem, at the level of neural systems that integrate functional regions of the telencephalon.

1 Introduction

Data from cognitive psychology, neurophysiology, and neuroanatomy indicate unequivocally that the properties of objects and events that we perceive through various sensory channels engage geographically separate sensory regions of the brain (Posner 1980; Van Essen and Maunsell 1983; Damasio 1985; Livingstone and Hubel 1988). The need to “bind” together the fragmentary representations of visual information has been noted by Treisman and Gelade (1980), Crick (1984), and others, but clearly the problem is a much broader one and includes the need to integrate both the sensory and motor components in both perception and recall, at all scales and at all levels. This broader concept of binding is closer to that of Sejnowski (1986).

The traditional and by now untenable solution to the binding problem has been that the components provided by different sensory portals end up being displayed together in so-called multimodal cortices, where the most detailed and integrated representations of reality are achieved. This intuitively reasonable view suggests that perception depends on a unidirectional process which provides a gradual refinement of signal extraction

along a cascade aimed towards integrative cortices in anterior temporal and anterior frontal regions. Some of the most influential accounts for the neural basis of cognition in the post-war period as well as major discoveries of neurophysiology and neuroanatomy over the past two decades, have seemed compatible with this view. After all, anatomical projections do radiate from primary sensory cortices toward structures in the hippocampus and prefrontal cortices via a multi-stage sequence (Pandya and Kuypers 1969; Jones and Powell 1970; Nauta 1971; Van Hoesen 1982), and the farther away neurons are from primary sensory cortices, the larger their receptive fields become, and the less unimodal their responses are (Desimone and Ungerleider 1989). However, there are several lines of evidence on the basis of which this traditional solution can be rejected.

2 Experimental Evidence

Evidence from Experimental Neuroanatomy: The notion that integration of perceptual or recalled components depends on a single neural meeting ground calls for the identification of a neuroanatomical site that would receive projections from all neural regions involved in the processing of entities and events as they occur in experience. Despite considerable exploration no such region has yet been found. The anterior temporal cortices and the hippocampus do receive projections from multiple sensory areas, but not from motor regions (Van Hoesen 1982). The anterior frontal cortices, the most frequently mentioned candidates for ultimate integration, are even less suited for that role. The sensory and motor streams that reach them remain segregated in different regions (Goldman-Rakic 1988). In other words, there seems to be no structural foundation to support the intuition that temporal and spatial integration occur at a single site.

Advances in experimental neuroanatomy have added a new element to neuroanatomical reasoning about this problem: at every stage of the chain of forward cortical projections, there exist prominent projections back to the originating sites. Moreover, the systems are just as rich in multi-stage, reciprocating feedback projections as they are in feedforward projections (Van Hoesen 1982; Van Essen 1985; Livingstone and Hubel 1987). The neuroanatomical networks revealed by these studies allow for both forward convergence of some parallel processing streams, and for the flow of signaling back to points of origin. In the proposal we will describe below, such networks operate as coherent phase-locked loops in which patterns of neural activity in "higher" areas can trigger, enhance, or suppress patterns of activity in "lower" areas.

Evidence from Experimental Neuropsychology in Humans with Focal Cerebral Lesions: If temporal and frontal integrative cortices were the substrate for the integration of neural activity on which binding depends, the bilateral destruction of those cortices in humans should: (a) preclude the

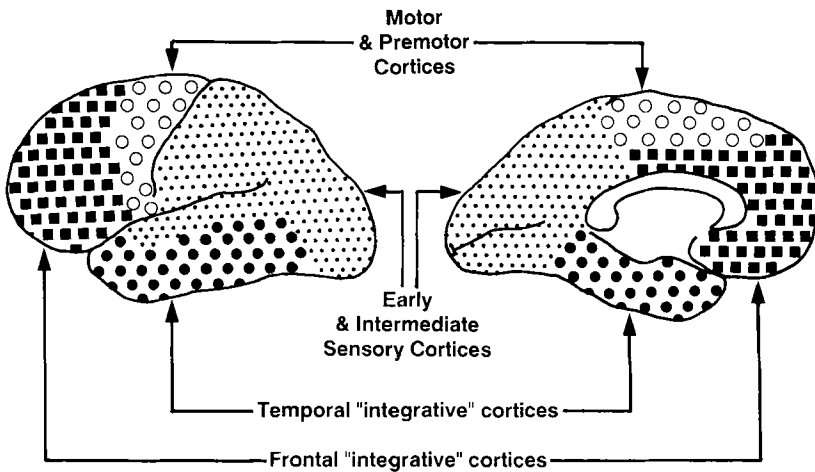


Figure 1: Fundamental divisions of human cerebral cortex depicted in a simplified diagram of the external and internal views of the left hemisphere. The *motor and premotor cortices* include cytoarchitectonic fields 4, 6, and 8. The *early and intermediate sensory cortices* include the primary visual, auditory, and somatosensory regions (respectively fields 17, 41/42, and 3/1/2), and the surrounding association cortices (fields 18/19, 7, 39, 22, 40, 5). The *temporal "integrative" cortices* include fields 37, 36, 35, 20, 21, 38, and 28, i.e., neocortical as well as limbic and paralimbic areas. The *frontal "integrative" cortices* include fields 44, 45, 46, 9, 10, 11, 12, 13, and 25, i.e., prefrontal neocortices as well as limbic.

perception of reality as a coherent multimodal experience and reduce experience to disjointed, modal tracks of sensory or motor processing; (b) reduce the integration of even such modal track processing; and (c) disable memory for any form of past integrated experience and interfere with all levels and types of memory. However, the results of bilateral destruction of the anterior temporal lobes, as well as bilateral destruction of prefrontal cortices, falsify these predictions (see Fig. 1).

Coherent perceptual experience is not altered by bilateral damage to the anterior temporal regions, nor does such damage disturb perceptual quality (see Corkin 1984; Damasio et al. 1985; Damasio et al. 1987). Our patient Boswell is a case in point. His extensive, bilateral damage in

anterior temporal cortices and hippocampus, disables his memory for unique autobiographical events, but not his ability to perceive the world around in fully integrated fashion and to recall and recognize the entities and events that he encounters or participates in, at non-unique level. His binding ability breaks down at the level of unique events, when the integration of extremely complex combinatorial arrangement of entities is required.

Bilateral lesions in prefrontal cortices, especially when restricted to the orbitofrontal sector, are also compatible with normal perception and even with normal memory for most entities and events except for those that pertain to the domain of social knowledge (Eslinger and Damasio 1985; Damasio and Tranel 1988).

Finally, it is damage to certain sectors of sensory association cortices that can affect both the quality of some aspects of perception within the modality served by those cortices, and recognition and recall. Depending on precisely which region of visual cortex is affected, lesions in early visual association cortices can disrupt perception of shape, or color, or texture, or stereopsis, or spatial placement of the physical components of a stimulus (Damasio 1985; Damasio et al. 1989). A patient may lose the ability to perceive color and yet perceive shape, depth and motion normally. More importantly, 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. For instance, patients may become unable to recognize familiar faces that they perceive flawlessly (although, intriguingly, they can discriminate familiar from unfamiliar faces at covert level; Tranel and Damasio 1985; 1988).

The key point is that damage in a posterior and unimodal association cortex can disrupt recall and recognition at virtually every level of the binding chain, from the entity-categorical level to the event-unique level. It can preclude the kind of integrated experience usually attributed to the anterior cortices.

3 A New View on the Binding Problem

The evidence then indicates: (a) that substantial binding, relative to entities or parts thereof, occurs in unimodal cortices and can support recall and recognition at the level of categories; (b) that recall and recognition at category level, are generally not impaired by damage confined to anterior integrative cortices, i.e., knowledge recalled at categoric levels depends largely on posterior sensory cortices and interconnected motor cortices; (c) that recall and recognition of knowledge at the level of unique entities or events, requires both anterior and posterior sensory cortices, i.e., a more complex network is needed to map uniqueness; anterior integra-

tive structures alone are not sufficient to record and reconstruct unique knowledge.

The implication is that the early and intermediate posterior sensory cortices contain fragmentary records of featural components which can be reactivated, on the basis of appropriate combinatorial arrangements (by fragmentary featural components we mean “parts of entities,” at a multiplicity of scales, most notably at feature level, e.g., color, movement, texture, shape and parts thereof). They also contain records of the combinatorial arrangement of features that defined entities (“local” or “entity” binding), but do not contain records of the spatial and temporal relationships assumed by varied entities within an event (“non-local” or “event binding”). The latter records, the complex combinatorial codes needed for event recall, are inscribed in anterior cortices. In this perspective the posterior cortices contain the fragments with which any experience of entities or events can potentially be re-enacted, but only contain the binding mechanism to re-enact knowledge relative to entities. Posterior cortices require binding mechanisms in anterior structures in order to guide the pattern of multiregional activations necessary to reconstitute an event.

Thus posterior cortices contain both basic fragments and local binding records and are essential for recreating any past experience. Anterior cortices contain non-local or event-binding records and are only crucial for reconstitution of contextually more complex events.

Perhaps the most important distinction between this perspective and the traditional view, is that higher-order anterior cortices are seen as repositories of combinatorial codes for inscriptions that lie elsewhere and can be reconstructed elsewhere, rather than being the storage site for the more refined “multimodal” representations of experiences. Although anterior cortices receive multimodal projections we conceptualize the records they harbor as amodal.

If parts of the representation of an entity are distributed over distant regions of the brain, then mechanisms must be available to bind together the fragments. A proposal for a new solution to the binding problem (Damasio 1989) is illustrated in figure 2 and presented in outline as follows:

1. The neural activity prompted by perceiving the various physical properties of any entity, occurs in fragmented fashion and in geographically separate regions located in early sensory cortices and in motor cortices. So-called “integrative” cortices do not contain such fragmentary inscriptions.
2. The integration of multiple aspects of external and internal reality in perceptual or recalled experiences, depends on the phase-locked co-activation of geographically separate sites of neural activity within the above mentioned sensory and motor cortices, rather than on a transfer and spatial integration of different representations towards anterior higher-order cortices. Consciousness of those co-

activations depends on their being attended to, i.e., on simultaneous enhancement of a pertinent set of activity against background activity.

3. The patterns of neural activity that correspond to distinct physical properties of entities are recorded in the same neural ensembles

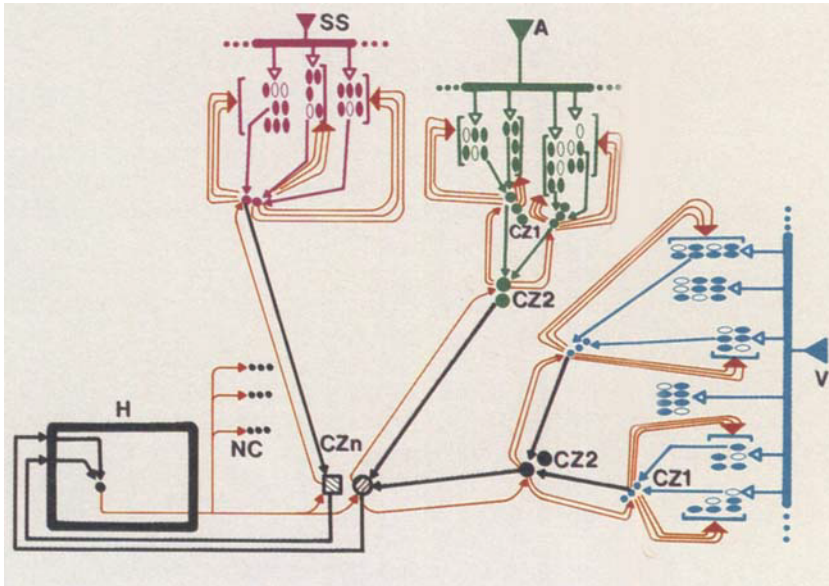


Figure 2: Simplified diagram of some aspects of the proposed neural architecture. *V*, *SS*, and *A* depict early and intermediate sensory cortices in visual, somatosensory, and auditory modalities. In each of those sensory sectors, separate functional regions are represented by open and filled dots. Note feedforward projections (black lines) from those regions toward several orders of convergence zones (*CZ1*, *CZ2*, *CZn*), and note also feedback projections from each *CZ* level toward originating regions (red lines). *H* depicts hippocampal system, one of the structures where signals related to a large number of activity sites can converge. Note outputs of *H* toward last station of feedforward convergence zones (*CZn*) and toward noncortical neural stations (*NC*) in basal forebrain, brain stem, and neurotransmitter nuclei. Feedforward and feedback pathways should not be seen as rigid channels. They are conceived as facilitated lines which become active when concurrent firing in early cortices or *CZs* takes place. Furthermore, those pathways terminate over neuron ensembles, in distributed fashion, rather than on specific single neurons.

in which they occur during perception, but the combinatorial arrangements (binding codes) that describe their pertinent linkages in entities and in events (their spatial and temporal coincidences), are stored in separate neural ensembles called convergence zones.

4. Convergence zones trigger and synchronize neural activity patterns corresponding to topographically organized fragment representations of physical structure, that were pertinently associated in experience, on the basis of similarity, spatial placement, temporal sequence, or temporal coincidence, or combinations thereof. The triggering and synchronization depends on feedback projections from the convergence zone to multiple cortical regions where fragment records can be activated.
5. Convergence zones are located throughout the telencephalon, at multiple neural levels, in association cortices of different orders, limbic cortices and subcortical limbic nuclei, and non-limbic subcortical nuclei such as the basal ganglia.
6. The geographic location of convergence zones for different entities varies among individuals but is not random. It is constrained by the subject matter of the recorded material (its domain), and by contextual complexity of events (the number of component entities that interact in an event and the relations they adopt), and by the anatomical design of the system. Convergence zones that bind features into entities are located earlier in the processing streams, and convergence zones that bind entities into progressively more complex events are gradually placed more anteriorly in the processing streams.
7. The representations inscribed in the above architecture, both those that preserve topographic/topologic relationships and those that code for combinatorial arrangements, are committed to populations of neuron ensembles and their synapses, in distributed form.
8. The co-occurrence of activities in multiple sites that is necessary for binding conjunctions, is achieved by recurrent feedback interactions.

Thus, we propose that the processing does not proceed in a single direction but rather through temporally coherent phase-locking amongst multiple regions. Although the convergence zones that realize the more encompassing integration are placed more anteriorly, it is activity in the more posterior cortical regions that is more directly related to conscious experience.

By means of feedback, convergence zones repeatedly return processing to earlier cortices where activity can proceed again towards the same or other convergence zones. Integration takes place when activations occur within the same time window, in earlier cortices. There is no need

to postulate a "final" and single integration area. This model accommodates the segregation of neural processing streams that neuroanatomical and neurophysiological data continue to reveal so consistently, and is compatible with the increase in receptive fields of neurons that occurs in cerebral cortex, in the posterior-anterior direction. It accords with the proposal 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. Broad receptive field neurons serve as pivots for reciprocating feedback projections rather than as accumulators of the knowledge inscribed at earlier levels. They are intermediaries in a continuous process that centers on early cortices.

4 Conclusions

The problem of how the brain copes with the fragmentary representations of information is central to our understanding of brain function. It is not enough for the brain to analyze the world into its components parts: the brain must bind together those parts that make whole entities and events, both for recognition and recall. Consciousness must necessarily be based on the mechanisms that perform the binding. The hypothesis suggested here is that the binding occurs in multiple regions that are linked together through activation zones; that these regions communicate through feedback pathways to earlier stages of cortical processing where the parts are represented; and that the neural correlates of consciousness should be sought in the phase-locked signals that are used to communicate between these activation zones.

Several questions are raised by this new view. For instance, what is the precise nature of the feedback signals that provide temporally coherent phase-locking among multiple regions? How large are the convergence zones in different parts of the brain? How are the decisions made to store an aspect of experience in a particular zone? There are several possible approaches to test the hypothesis proposed here. One approach is to develop new techniques for recording from many neurons simultaneously in communicating brain regions. Another relies on neuropsychological experiments in neurological patients with small focal lesions in key areas of putative networks dedicated to specific cognitive processes. Finally, modeling studies should illuminate the collective properties of convergence zones and provide us with the intuition we need to sharpen our questions.

Acknowledgments

Supported by NINCDS Grant PO1 NS19632.

References

- 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. *Proc. Natl. Acad. Sci. USA* 81, 4586–4590.
- Damasio, A. 1989. Multiregional Retroactivation: A Systems Level Model for Some Neural Substrates of Cognition. *Cognition*, in press.
- . 1985. Disorders of Complex Visual Processing. In: *Principles of Behavioral Neurology*, ed. M.M. Mesulam, Contemporary Neurology Series, 259–288. Philadelphia: F.A. Davis.
- Damasio, A., P. Eslinger, H. Damasio, G.W. Van Hoesen, and S. Cornell. 1985. Multimodal Amnesic Syndrome Following Bilateral Temporal and Basal Forebrain Damage. *Archives of Neurology* 42, 252–259.
- Damasio, A., H. Damasio, D. Tranel, K. Welsh, and J. Brandt. 1987. Additional Neural and Cognitive Evidence in Patient DRB. *Society for Neuroscience* 13, 1452.
- Damasio, A. and D. Tranel. 1988. Domain-specific Amnesia for Social Knowledge. *Society for Neuroscience* 14, 1289.
- Damasio, A.R., H. Damasio, and D. Tranel. 1989. Impairments of Visual Recognition as Clues to the Processes of Memory. In: *Signal and Sense: Local and Global Order in Perceptual Maps*, eds. G. Edelman, E. Gall, and M. Cowan, Neuroscience Institute Monograph. Wiley and Sons.
- Desimone, R. and L. Ungerleider. 1989. Neural Mechanisms of Visual Processing in Monkeys. In: *Handbook of Neuropsychology, Disorders of Visual Processing*, ed. A. Damasio, in press.
- Eslinger, P. and A. Damasio. 1985. Severe Disturbance of Higher Cognition after Bilateral Frontal Lobe Ablation. *Neurology* 35, 1731–1741.
- Goldman-Rakic, P.S. 1988. Topography of Cognition: Parallel Distributed Networks in Primate Association Cortex. In: *Annual Review of Neuroscience II*, Annual Reviews Inc., Palo Alto, CA, 137–156.
- Jones, E.G. and T.P.S. Powell. 1970. An Anatomical Study of Converging Sensory Pathways within the Cerebral Cortex of the Monkey. *Brain* 93, 793–820.
- Livingstone, M. and D. Hubel. 1988. Segregation of Form, Color, Movement, and Depth: Anatomy, Physiology, and Perception. *Science* 240, 740–749.
- . 1987. 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.
- Nauta, W.J.H. 1971. The Problem of the Frontal Lobe: A Reinterpretation. *J. Psychiat. Res.* 8, 167–187.
- Pandya, D.N. and H.G.J.M. Kuypers. 1969. Cortico-cortical Connections in the Rhesus Monkey. *Brain Res.* 13, 13–36.
- Posner, M.I. 1980. Orienting of Attention. *Quarterly Journal of Experimental Psychology* 32, 3–25.
- Sejnowski, T.J. 1986. Open Questions about Computation in Cerebral Cortex. In:

- Parallel Distributed Processing, eds. J.L. McClelland and D.E. Rummelhart, 372–389. Cambridge: MIT Press.
- Tranel, D. and A. Damasio. 1988. Nonconscious Face Recognition in Patients with Face Agnosia. *Behavioral Brain Research* 30, 235–249.
- . 1985. Knowledge without Awareness: An Autonomic Index of Facial Recognition by Prosopagnosics. *Science* 228:21, 1453–1454.
- Treisman, A. and G. Gelade. 1980. A Feature-integration Theory of Attention. *Cognitive Neuropsychology* 12, 97–136.
- Van Essen, D.C. 1985. Functional Organization of Primate Visual Cortex. In: Cerebral Cortex, eds. A. Peters and E.G. Jones, 259–329. Plenum Publishing.
- Van Essen, D.C. and J.H.R. Maunsell. 1983. Hierarchical Organization and Functional Streams in the Visual Cortex. *Trends in Neuroscience* 6, 370–375.
- Van Hoesen, G.W. 1982. The Primate Parahippocampal Gyrus: New Insights Regarding its Cortical Connections. *Trends in Neurosciences* 5, 345–350.

Received 18 November; accepted 25 November 1988.