Neurological disorders and the structure of human consciousness

Jeffrey W. Cooney[1] and Michael S. Gazzaniga[2]

[1]Department of Psychology, Dartmouth College, 6162 Moore Hall, Hanover, NH 03755, USA;
currently at: Helen Wills Neuroscience Institute, University of California, Berkeley, 5201 Tolman Hall, Berkeley, CA 94720, USA
[2]Center for Cognitive Neuroscience and Department of Psychology, Dartmouth College, 6162 Moore Hall, Hanover, NH 03755, USA

Recent studies that identify distinct neural correlates of perceptual awareness offer a promising step towards improved understanding of the neurological underpinnings of conscious experience. Such studies indicate that perceptual awareness is modular in nature, with neural correlates of awareness consisting of the specialized structures involved in perceptual processing. However, the integrative, multimodal nature of conscious experience appears to require a functional architecture that overcomes this modular segregation of function. We propose a model in which experience emerges from the dynamic interactions of specialized component processes via a distributed neural network. Such a model offers a mechanism to explain several empirical observations of the neural correlates of perceptual awareness, cognitive function, and symptoms of neurological damage.

The conclusions of an interpretive system are only as good as the information the system receives. Applied to the human brain, this principle might help to make sense of an array of peculiar neurological conditions that involve the formation of either incomplete or delusional understandings about oneself, other objects or individuals, and the surrounding environment. Such beliefs are central to the conditions of, for example, anosognosia for hemiplegia, hemispatial neglect and reduplicative paramnesia. In these conditions, key aspects of patients' conscious experience seem to be either altered or to disappear altogether: paralysis is denied, left space ignored, and locations duplicated and confused. Strikingly, patients with these syndromes can be profoundly unaware of their limitations, developing and maintaining delusional beliefs regardless of the evidence to the contrary, even to the point of denying ownership of their own limbs [1].

We argue that a range of bizarre pathologies may be understood as arising from a modular organization of the brain in which conscious experience comprises components that directly reflect and result from the modular nature of neural circuits specialized to process specific domains of information [2–6]. Each specialized neural circuit enables the processing and mental representation of specific aspects of conscious experience, with the outputs of the distributed circuits integrated via a large-scale neuronal 'workspace' [7–9]. We propose that, when combined with an interpretive process that occurs in the left hemisphere of the human brain [2,10,11], a modular organization of the information processing that underlies conscious experience can explain a variety of neurological syndromes. In this model, such syndromes result from the attempts of an interpretive system to construct a coherent story after the ability to process specific aspects of conscious experience has been either altered or abolished as a result of neural injury.

Modularity of perceptual awareness

Neurophysiological evidence for the modular organization of perceptual awareness is strong. Rather than functioning as a general-purpose computing device, the brain consists of a network of adaptive systems [2,11]. Functional neuroimaging experiments, single-cell recordings and neuropsychological dissociations of function indicate that neural circuits segregate into functionally specialized systems on both cognitive and cellular levels, and reveal a robust correlation between the performance of specific cognitive tasks and the activation of specific regions of neural tissue [4,5,9,12,13].

For example, regions such as the Fusiform Face Area (FFA), Parahippocampal Place Area (PPA), and area V5 of the visual cortex are activated selectively, not only by exposure to face, place and moving stimuli, respectively, but by mental imagery of such stimuli as well [12,13]. More strikingly, the activity of these regions is observed to vary directly with the contents of awareness during binocular rivalry tasks, increasing markedly when the stimulus to which each region preferentially responds is consciously perceived [4,5]. In such experiments, the presentation of a separate stimulus to each eye results in an alternating conscious perception of first one stimulus, then the other, rather than a blended image or the perception of both stimuli simultaneously. Because the presented stimuli remain constant throughout the task, the co-variation of neural activation with the dominant percept represents a direct correlation between neural activity and conscious experience. In addition, the finding that direct electrical stimulation of specific brain regions can elicit perceptual experiences related to the processing role of the stimulated circuit [4,14] strengthens the modular-consciousness hypothesis. These data indicate that activation of specific neural

Corresponding author: Jeffrey W. Cooney (jcooney@uclink.berkeley.edu).

http://tics.trends.com 1364-6613/03/$ - see front matter doi:10.1016/S1364-6613(03)00058-5

regions might not only correlate with specific perceptual experiences, but could be sufficient to cause them.

Integrated awareness emerges from modular interactions within a neuronal workspace

Although evidence for the modularity of information processing is strong, a modular architecture is not, in itself, sufficient to account for the majority of cognitive tasks that occur within the realm of conscious experience. The presence of a large-scale network, whose long-range connectivity provides a neural workspace through which the outputs of numerous, specialized, brain regions can be interconnected and integrated, provides a promising solution to this need for integration, and fits well with a range of empirical observations outlined in the following sections. Although the characteristics of such a network are not clear currently, several studies indicate the involvement of prefrontal cortex, anterior cingulate [9,15] and thalamocortical interactions [16] in mechanisms that enable the mobilization of modular information into a state of large-scale network availability [17].

In the workspace model, outputs from an array of parallel processors continually compete for influence within the network [7–9]. Dynamic integration of these outputs, which combines factors including bottom-up stimulus attributes, modulation caused by contextual valence and memory, and selective attentional amplification, determines which aspects of the available information emerge as dominant, and gives rise to a coherent network state in which the integrated information is widely available and can be used by a variety of cognitive processes that require input from multiple modules [9]. The result of this process is a fluctuating stream of transiently self-sustained, self-modifying workspace states, the characteristics of which are postulated to determine the contents of the subjective experience of the individual [9,18]. Although the majority of cerebral processing takes place outside conscious awareness, information is perceived consciously when it is sufficiently amplified within the system to generate a coherent network state in which the information is widely available to a range of modular processors. The population of neurons activated at any given moment then further influences the development of subsequent states of the network and facilitates a coherent progression of cognitive processes with no need for a separate, higher-order executive system.

The left hemisphere plays a unique role in shaping the contents of subjective experience

Studies of split-brain patients reveal that the left hemisphere of the human brain has a unique capacity to reflexively formulate causal theories about why events occur [2,10,11]. The adaptive benefit of such a process is immense, enabling information from diverse events to be combined into a chain of causal understanding that can guide future behavior. Although the mechanism through which the interpretive process works is not well understood, it appears to be an inherent property of the interactions of components of the system, and its function in shaping the contents of conscious experience is significant. Strikingly, studies of both split-brain and neurologically normal individuals indicate that the interpretive process continues to function when the range of available information is incomplete. This generates a causal understanding of events that is subjectively complete and seemingly self-evident, even when that understanding is incorrect [10,11,19] (Fig. 1). In the workspace model, this result could arise from the integration of information from an incomplete set of data. The process has evolved to provide a causal explanation and does so, but when the range of information available to the workspace is limited because of damage to the neural

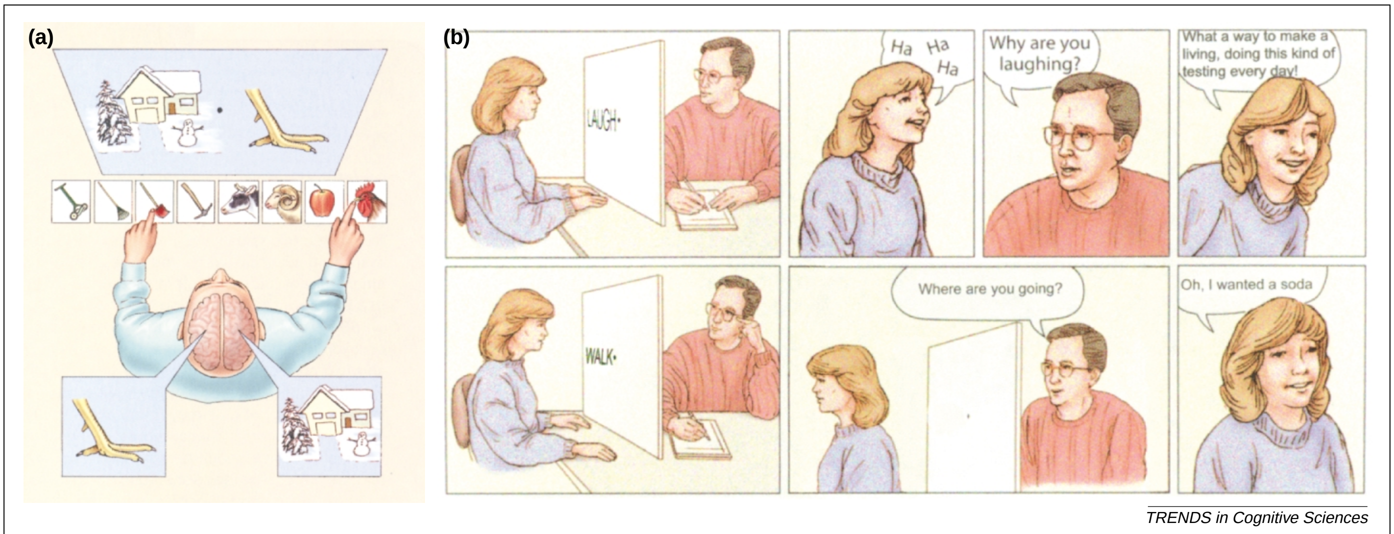

Fig. 1. Studies of split-brain patients reveal that the left hemisphere automatically constructs a subjective understanding of information available to it. (a) In a classic example of this process, a picture of a chicken claw was briefly presented to the (speaking) left hemisphere, while a snow scene was concurrently shown to the (silent) right hemisphere. After these initial images were removed, patient PS easily picked out related pictures from a set of eight choices, which remained on the screen. His left hand pointed to a snow shovel and his right hand pointed to a chicken, and both hemispheres were able to observe these choices. When asked why he had picked those particular pictures, PS said '*Oh, that's simple. The chicken claw goes with the chicken, and you need a shovel to clean out the chicken shed.*' (Reproduced with permission from Ref. [26].) (b) Presentation of a command to the non-speaking right hemisphere generates behavior based upon information which is unavailable to the left hemisphere. After the action is performed, however, the speaking left hemisphere automatically generates an understanding of the behavior (two examples given). Although the understanding is incorrect, it is experienced by that hemisphere as an accurate description of the event. (Adapted with permission from Ref. [26]).

processors that provide information to the network, the accuracy of the resulting interpretation is limited to the restricted range of available information.

Interpreting a post-lesion world

When considered within a framework of modular, interpreted consciousness, a variety of seemingly bizarre neurological syndromes reveal a rational pattern of symptoms. Although phenomenally distinct, such syndromes may be mechanistically related, with symptoms falling into place as the logical results of an interpretive system that is striving to make sense of an altered set of available information.

Neurological lesions can affect perception and memory

Because particular lesions leave patients with no experience of deficit, it follows that these patients must have a satisfactory agreement between their post-lesion experience and their pre-lesion memories. Such an extension of neurological symptoms into the realm of memory has been demonstrated in patients with hemispatial neglect, whose left-neglected memories match their left-neglected experience of the world [20]. In the modular-consciousness model, this result is not surprising. Because the neural correlates of consciousness are proposed to consist of activity in the very networks that are responsible for processing the relevant informational attributes, this model predicts that focal lesions might have dramatic effects on specific aspects of conscious experience, including their representation in memory. Just as particular capacities, such as face recognition, generation of nouns in speech and manipulation of objects in space might be either impaired or eliminated by cortical damage, so too might the ability to process and represent specific aspects of conscious experience be either impaired or abolished by neurological injury.

Studies indicating that retrieval of a memory entails activation of the same perceptual circuits that are activated directly by the event being recalled [12,21] complete the picture. For patients with visual hemineglect, neglect of the left visual field occurs in the realm of memory, just as it does during actual perception of a scene, although information about the complete spatial scene remains encoded in the brain [20] (Fig. 2). The completed memory cannot be reconstructed from any one mental vantage point because, from each perspective, the aspects of the memory that pertain to left visual space cannot receive sufficient attentional amplification to enable their large-scale representation and availability in the neuronal workspace, and, thus, in conscious experience. Changing one's perspective in the visual image enables the information neglected from one vantage point to become available from the second. This is because the change in egocentric coordinates enables information about the previously neglected space to be processed by neural circuits that are responsible for handling information about other parts of visual space that remain able to receive the necessary attentional amplification to bring the stored information into consciousness.

The patient's left-neglected memories therefore directly parallel his left-neglected perception of the surrounding

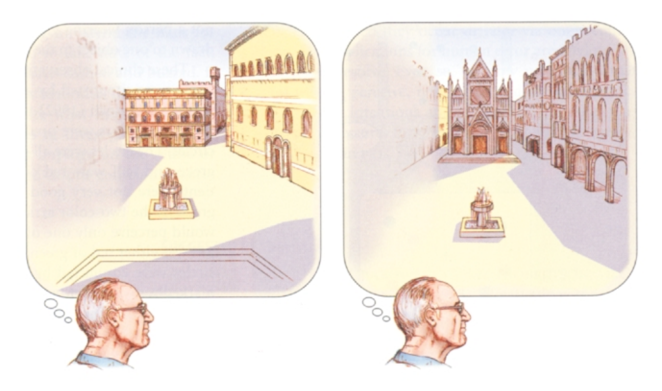

Fig. 2. In an experiment designed to assess the effects of hemispatial neglect on memory representations [18], patients with left hemineglect were asked to visualize themselves in a piazza familiar to them before the onset of neglect, and to describe the scene from the opposite ends of the piazza. The descriptions provided from each imagined position neglected features of the contralesional side of the scene, indicating that although the entire memory was preserved, the aspects of the memory that could be brought into conscious awareness varied systematically with the imagined perspective of the patient. Therefore, if the patient were actually to visit the piazza, there would be no difference in the comparison of his pre-lesion memories with his post-lesion perceptual experience, because his neglect affects conscious experience of both the memory and the perception. (Reproduced with permission from Ref. [26]).

world because both result from damage to the same process of attention allocation in the workspace network. Looking around a familiar scene, everything seems just as it should, because the memory of the scene is itself left-neglected. If a particular landmark is not observed, the neglect patient might rotate until it comes into view [22], just as a neurologically normal individual might rotate to find something located behind him. In each case, the individual draws upon his memory's representation of the scene and uses this to navigate through his conception of space. The patient experiences no alteration in perception because memory and perception remain in agreement, with no discrepancy to indicate a problem.

Focal lesions can abolish the ability to process entire realms of conceptual information

Loss of the ability to process information about specific aspects of conscious experience might also account for the inability of some patients to learn of their impairment through other cognitive means. This inability to use semantic knowledge to analyze their condition remains puzzling. However, the finding that some patients with anosognosia for hemiplegia are unable to identify paralysis in themselves as well as in others [23] indicates that they might have lost the ability to process information accurately from the entire domain in which their impairment exists. Their brains are impaired in their ability to instate workspace states that represent left-body information about themselves and others, and so such information remains excluded from their available range of conscious experience. Indicative of this redefinition of the range of conscious experience, one patient explained that '*I knew the word "neglect" was a sort of medical term for whatever was wrong but the word bothered me because you only neglect something that is actually there, don't you? If it's not there, how can you neglect it?*' [24]. With no memories to indicate a deficit, and with an impaired ability to process information that is related to the experiential realm in which the deficit exists, it begins to be more understandable that some patients can be unaware of their impairments.

Denials of deficit result from interpretations of incomplete information

In the proposed model, the denials of deficit observed in visual neglect and anosognosia for hemiplegia can also be understood as the logical results of damage to specific sets of neural circuits. Whereas paralysis due to severing of nerves that connect the parietal cortex to the body would be noted immediately, damage to the (right) parietal cortex itself leaves no part of the brain to monitor or represent the existence of limbs on the left side of the body. The region responsible for signaling a problem has, itself, been damaged, leaving no system to file a complaint. Therefore, the patient no longer registers the existence of left limbs but has an intact interpretive network that concludes that everything must be fine because no problem has been indicated internally. Furthermore, because the workspace network contains no higher-order system, the absence of information about the left side of one's body is no more worrisome than a lack of visual information from behind one's head – no impairment is registered because no input is expected. Therefore, when unable to do a bimanual task, the system generates a reasonable story, such as '*I didn't want to do that task*.' [1]. Presented with a paralyzed left hand, and accompanied by neither sensory input nor registration of the lack thereof, the system draws a subjectively reasonable conclusion: '*That's not my hand*' [1]. Such attempts at rationalization mimic the accounts put forth by the speaking left hemispheres of split-brain patients following right-hemisphere-generated behavior, and are merely extensions of the interpretive processes that allow us to make sense of the world around us on an everyday basis.

Last, the delusions of reduplicative paramnesia are prime examples of an interpretive system at work with a corrupted set of information. Patients with this syndrome appear to have abnormal neural indicators of familiarity and believe that multiple copies of places and people exist [25]. While being examined at New York Hospital, one patient was convinced that she was in her own home in Freeport, Maine. Although intelligent and aware that her doctors disagreed with her assessment, she remained adamant in her assertions. When pressed to explain why, if she were indeed in her house in Freeport, there was a bank of elevators outside the door, she replied calmly '*Doctor, do you know how much it cost me to have those put in?*' [11]. Such patients are strikingly unflappable, incorporating incongruous information with ease. The woman in this example 'knew' that she was in her own house, just as we know when we are in ours, and the interpretive processes of the workspace network makes such information seem self-evident, even when it is wrong. However, although the ability of individuals with delusional syndromes to create a subjectively rational story out of seemingly incoherent information is startling, what might be most striking is the realization that such syndromes reveal the overwhelming significance of this process in shaping our own experience of the world.

The explanations provided by an interpretive system are only as good as the information available to it, and bizarre information yields bizarre results. Patients with delusional syndromes might know that their assertions sound strange to others, but they also 'know' that the scene they perceive is complete, that the hand presented to them is not theirs, and that they are sitting in their own homes. The information presented by the modular components of the workspace network are all that the brain has available to it, and yields subjective experiences that are just as self-evident and unassailable as the experience engendered by any other network state. The presence of such an interpretive mechanism is beneficial in an intact brain and on an evolutionary scale because it greatly enhances the ability of an individual to adapt to a wide range of novel and unexpected situations. However, it results in a variety of strange beliefs when the internal information of the brain is itself corrupted.

Conclusions

Data from a wide range of functional, physiological and clinical studies support a model of consciousness in which subjective awareness emerges from the interactions of specialized, modular components in a distributed neural network. In addition, the activity of such components is united cognitively by an interpretive process that occurs in the left hemisphere of the human brain. The dynamic, self-modifying nature of the interactions between these components is driven solely by the range of information available to the system and enables the coherent progression of cognitive processes without requiring a higher-order executive system. Furthermore, the specialized nature of the component neural processors suggests a potential explanation for how specific lesions can selectively alter or eliminate the contents of particular aspects of subjective experience, with no resulting experience of deficit for the brain-damaged individual.

Although the proposed model remains in development as a conceptual framework, we believe that, for the available data, a model based on the interpreted integration of specialized neural processors has notable explanatory power. Further study is needed to better characterize the nature of these modular processors and the ways in which they interact. We believe that such work will provide significant insight into the processes that underlie the construction of subjective experience.

References

1 Ramachandran, V.S. (1995) Anosognosia in parietal lobe syndrome. *Conscious. Cogn.* 4, 22–51
2 Gazzaniga, M.S. (1989) Organization of the human brain. *Science* 245, 947–952
3 Kanwisher, N. (2001) Neural events and perceptual awareness. *Cognition* 79, 89–113
4 Logothetis, N. (1998) Single units and conscious vision. *Philos. Trans. R Soc Lond. B Biol. Sci.* 353, 1801–1818
5 Tong, F. *et al.* (1998) Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21, 753–759
6 Culham, J.C. *et al.* (1999) Recovery of fMRI activation in motion area MT following storage of the motion aftereffect. *J. Neurophysiol.* 81, 388–393

7 Baars, B.J. (1989) *A Cognitive Theory of Consciousness*, Cambridge University Press

8 Dehaene, S. (1998) A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14529–14534

9 Dehaene, S. and Naccache, L. (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37

10 Gazzaniga, M.S. and LeDoux, J.E. (1978) *The Integrated Mind*, Plenum Press

11 Gazzaniga, M.S. (2000) Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain* 123, 1293–1326

12 O'Craven, K. and Kanwisher, N. (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023

13 Kanwisher, N. *et al.* (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311

14 Puce, A. *et al.* (1999) Electrophysiological of human face perception: III. Effects of top-down processing on face-specific potentials. *Cereb. Cortex* 9, 445–458

15 Posner, M.I. (1994) Attention: the mechanisms of consciousness. *Proc. Natl. Acad. Sci. U. S. A.* 91, 7398–7403

16 Newman, J. and Baars, B.J. (1993) A neural attentional model for access to consciousness: a global workspace perspective. *Concepts Neurosci.* 4, 255–290

17 Baars, B. (2001) The conscious access hypothesis: origins and recent evidence. *Trends Cogn. Sci.* 6, 47–52

18 Dennet, D.C. (2001) Are we explaining consciousness yet? *Cognition* 79, 221–237

19 Schacter, D. and Singer, J.E. (1962) Cognitive, social, and physiological determinants of emotional state. *Psychol. Rev.* 69, 379–399

20 Bisiach, E. and Luzzatti, C. (1978) Unilateral neglect of representation space. *Cortex* 14, 129–133

21 Kosslyn, S.M. *et al.* (2001) Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642

22 Sacks, O. (1970) Eyes right! *The Man Who Mistook His Wife for a Hat, and Other Clinical Tales*, pp. 77–79, Harper Perennial

23 Ramachandran, V.S. and Rogers-Ramachandran, D. (1996) Denial of disabilities in anosognosia. *Nature* 382, 501–502

24 Halligan, P.W. and Marshall, J.C. (1998) Neglect of awareness. *Conscious. Cogn.* 7, 356–380

25 Breen, N. *et al.* (2000) Towards an understanding of delusions of misidentification: four case studies. *Mind Lang.* 15, 74–110

26 Gazzaniga, M.S. *et al.*, eds. (1998) *Cognitive Neuroscience: The Biology of the Mind* W.W. Norton & Co