items from memory, and categorization processes generally. For example, the pronunciation model of Plaut et al. (1996) acquires a combinatorially structured set of output attractors encoding phonological strings including monosyllabic English words, and an input encoding a letter string yields an output activation pattern that is an attractor for a corresponding pronunciation.

A related principle governing processing in a class of recurrent networks characterizes the output of the network as an *optimal* activation pattern: among those patterns containing the given input pattern, the output is the pattern that maximizes a numerical well-formedness measure, harmony, or that minimizes "energy" (see also CONSTRAINT SATISFACTION). This principle has been used in combination with the following one to derive a general grammar formalism, harmonic grammar, described above as an illustration of principle-centered research. Harmonic grammar is a precursor to OPTIMALITY THEORY (Prince and Smolensky 1993), which adds further strong restrictions on what constitutes a possible human grammar. These include the universality of grammatical constraints, and the requirement that the strengths of the constraints be such as to entail "strict domination": the cost of violating one constraint can never be exceeded by any amount of violation of weaker constraints.

Connectionist Representational Principles

Research on the representational component of connectionist theory has focused on statistically based analyses of internal representations learned by networks trained on linguistic data, and on techniques for representing, in numerical activation patterns, information structured by linear precedence, attribute/value, and dominance relations (e.g., Smolensky 1990; see BINDING PROBLEM). While this research shows how complex linguistic representations may be realized, processed, and learned in connectionist networks, contributions to the theory of linguistic representation remain largely a future prospect.

See also COGNITIVE MODELING, CONNECTIONIST; DISTRIBUTED VS. LOCAL REPRESENTATION; NATIVISM

—*Paul Smolensky*

References

Elman, J. L. (1990). Finding structure in time. *Cognitive Science* 14: 179–211.

Elman, J., E. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. (1996). *Rethinking Innateness: A Connectionist Perspective on Development.* Cambridge, MA: MIT Press.

Jacobs, R. A., M. I. Jordan, and A. G. Barto. (1991). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science* 15: 219–250.

Legendre, G., Y. Miyata, and P. Smolensky. (1990). Harmonic grammar: A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society.* Cambridge, MA, pp. 388–395.

Pinker, S., and J. Mehler. (1988). *Connections and Symbols.* Cambridge, MA: MIT Press.

Plaut, D., and T. Shallice. (1994). *Connectionist Modelling in Cognitive Neuropsychology: A Case Study.* Hillsdale, NJ: Erlbaum.

Plaut, D. C., J. L. McClelland, M. S. Seidenberg, and K. Patterson. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 103: 56–115.

Prince, A., and P. Smolensky. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar.* RuCCS Technical Report 2, Rutgers Center for Cognitive Science, Rutgers University, Piscataway, NJ, and Department of Computer Science, University of Colorado at Boulder.

Rumelhart, D., and J. L. McClelland. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Vol. 2, *Psychological and Biological Models.* Cambridge, MA: MIT Press, pp. 216–271.

Seidenberg, M. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science* 275: 1599–1603.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artificial Intelligence* 46: 159–216.

Smolensky, P., M. C. Mozer, and D. E. Rumelhart, Eds. (1996). *Mathematical Perspectives on Neural Networks.* Mahwah, NJ: Erlbaum.

Touretzky, D. S., Ed. (1991). *Machine Learning* 7(2/3). Special issue on connectionist approaches to language learning.

Further Readings

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition* 48: 71–99.

Goldsmith, J., Ed. (1993). *The Last Phonological Rule: Reflections on Constraints and Derivations.* Chicago: University of Chicago Press.

Hare, M., and J. L. Elman. (1994). Learning and morphological change. *Cognition* 49.

Hinton, G. E. (1991). *Connectionist Symbol Processing.* Cambridge, MA: MIT Press.

Miikkulainen, R. (1993). *Subsymbolic Natural Language Processing: An Integrated Model of Scripts, Lexicon, and Memory.* Cambridge, MA: MIT Press.

Plunkett, K., and V. Marchman. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48: 21–69.

Sharkey, N., Ed. (1992). *Connectionist Natural Language Processing.* Dordrecht: Kluwer.

Wheeler, D. W., and D. S. Touretzky. (1993). A connectionist implementation of cognitive phonology. In J. Goldsmith, Ed., *The Last Phonological Rule: Reflections on Constraints and Derivations.* Chicago: University of Chicago Press, pp. 146–172.

Consciousness

Conscious mental states include sensations, such as the pleasure of relaxing in a hot bath or the discomfort of a hangover, perceptual experiences, such as the visual experience of a computer screen about half a meter in front of me, and occurrent thoughts, such as the sudden thought about how a problem can be solved. Consciousness is thus a pervasive feature of our mental lives, but it is also a perplexing one. This perplexity—the sense that there is something mysterious about consciousness despite our familiarity with sensation, perception, and thought—arises principally from the question of how consciousness can be the product of physical processes in our brains.

Ullin Place (1956) introduced a precursor of central state materialism for conscious states such as sensations. But the idea that types of conscious experience are to be identified with types of brain processes raises an important question, which can be made vivid by using Thomas Nagel's (1974) idea of WHAT-IT'S-LIKE to be in a certain state—and, more generally, the idea of there being something that it is like to be a certain creature or system. The question is, why should there be something that it is like for certain processes to be occurring in our brains? Nagel's famous example of what it is like to be a bat illustrates that our grasp of facts about the subjective character of experiences depends very much on our particular perceptual systems. Our grasp on physical or neurophysiological theories, in contrast, is not so dependent. Thus it may appear that subjective facts are not to be identified with the facts that are spelled out in those scientific theories. This Nagelian argument about the elusiveness of QUALIA is importantly similar to Frank Jackson's (1982, 1986) "knowledge argument" and similar responses have been offered to both (Churchland 1985, 1988; and for a reply, Braddon-Mitchell and Jackson 1996).

Ned Block's (1978) "absent qualia argument" is different from the arguments of Nagel and Jackson because it is specifically directed against FUNCTIONALISM: the idea that mental states are individuated by the causal roles they play in the total mental economy, rather than by the particular neurophysiological ways these roles are realized. The problem for functionalism is that we can imagine a system (e.g., Block's homunculi-headed system) in which there is nothing that it is like to be that system, even though there are, within the system, devices that play the various functional roles associated with sensations, perceptions, and thoughts. This argument is not intended for use against a physicalist who (in the style of Place and subsequent central state materialists) simply identifies conscious mental states with brain processes (pain with C-fibers firing, for example). The examples used in the absent qualia argument may, however, be used to support the claim that it is even logically possible there could be a physical duplicate of a normal human being that nevertheless lacked qualia (a "zombie"; Chalmers 1996).

It is a disputed question whether arguments like Nagel's can establish an ontological conclusion that consciousness involves something nonphysical (see MIND-BODY PROBLEM). But even if they cannot, there still appears to be a problem about consciousness; namely, it is a mystery why there should be something that it is like to undergo certain physical processes. This is what Joseph Levine (1983) has called the EXPLANATORY GAP. Jackson and Block both join Nagel in seeing a puzzle at this point, and Colin McGinn (1989) has argued that understanding how physical processes give rise to consciousness is cognitively beyond us (for a critical appraisal of McGinn's argument, see Flanagan 1992).

One possible strategy for demystifying the notion of consciousness is to claim that consciousness is a matter of thought about mental states. This is the "higher-order thought theory of consciousness" favored by David Rosenthal (1986). In this theory, consciousness, considered as a property of mental states, is analyzed in terms of consciousness *of* mental states, while consciousness *of* something is analyzed in terms of having a thought about that thing. Thus for a mental state to be a conscious mental state is for the subject of that state to have a thought about it. If the higher-order thought theory were to be correct, then the occurrence of consciousness in the physical world would not be any more mysterious than the occurrence of mental states, which are not in themselves conscious states, or the occurrence of thoughts about mental states.

However, there are some quite serious problems for the higher-order thought theory. One is that the theory seems to face a kind of dilemma. If the notion of thought employed is a demanding one, then there could be something that it is like for a creature to be in certain states even though the creature did not have (perhaps, even, could not have) any thoughts about those states. In that case, higher-order thought is not necessary for consciousness. But if the notion of thought that is employed is a thin and undemanding one, then higher-order thought is not sufficient for consciousness. Suppose, for example, that thought is said to require no more than having discriminative capacities. Then it seems clear that a creature, or other system, could be in a certain type of mental state, and could have a capacity to detect whether it was in a state of that type, even though there was nothing that it was like to be that creature or system.

More generally, work toward the demystification of consciousness has a negative and a positive aspect. The negative aspect consists in seeking to reveal unclarities and paradoxes in the notion of the subjective character of experience (e.g., Dennett 1988, 1991). The positive aspect consists in offering putative explanations of one or another property of conscious experience in neural terms. Paul Churchland (1988, 148) clearly illustrates how to explain certain structural features of our experiences of color (for example, that an experience of orange is more like an experience of red than it is like an experience of blue). The explanation appeals to the system of neural coding for colors that involves triples of activation values corresponding to the illumination reaching three families of cones, and to structural properties of the three-dimensional space in which they are plotted (see COLOR VISION). But while this is a satisfying explanation of those structural features of color experiences, it seems to leave us without any account of why it is like *anything at all* to see red. Why there are *any* experiential correlates of the neural codes is left as a brute unexplained fact. The demystifier of consciousness may then reply that this appearance of residual mystery is illusory, and that it is a product either of fallacies and confusions that surround the notion of the subjective character of experience or else of an illegitimately high standard imposed on explanation.

The notion of consciousness associated with the idea of the subjective character of experience, and which generates the "hard problem" of consciousness (Chalmers 1996), is sometimes called "phenomenal consciousness." There are several other notions for which the term *consciousness* is sometimes used (Allport 1988), including being awake, voluntary action, ATTENTION, monitoring of internal states, reportability, INTROSPECTION, and SELF-KNOWLEDGE. The distinctions among these notions are important, especially for the assessment of cognitive psychological and neuroscientific theories of consciousness (see CONSCIOUSNESS, NEUROBIOLOGY OF).

One particularly useful contrast is between phenomenal consciousness and "access consciousness" (Block 1995, 231): "A state is access-conscious if, in virtue of one's having the state, a representation of its content is (1) inferentially promiscuous, that is, poised to be used as a premise in reasoning, (2) poised for rational control of action, and (3) poised for rational control of speech. . . . [Access consciousness is] a cluster concept, in which (3)—roughly, reportability—is the element of the cluster with the smallest weight, though (3) is often the best practical guide to [access consciousness]." The two notions appear to be independent in the sense that it is possible to have phenomenal (P) consciousness without access (A) consciousness, and vice versa. An example of P-consciousness without A-consciousness would be a situation in which there is an audible noise to which we pay no attention because we are engrossed in conversation. As an example of A-consciousness without P-consciousness, Block (1995, 233) suggests an imaginary phenomenon of "superblindsight." In ordinary cases of BLINDSIGHT, patients are able to guess correctly whether there is, for example, an O or an X in the blind region of their visual field, even though they are unable to see either an O or an X there. The state that represents an O or an X is neither a P-conscious nor an A-conscious state. In superblindsight, there is still no P-consciousness, but now the patient is imagined to be able to make free use in reasoning of the information that there is an O, or that there is an X.

While the notion of phenomenal consciousness applies most naturally to sensations and perceptual experiences, the notion of access consciousness applies very clearly to thoughts. It is not obvious whether we should extend the notion of phenomenal consciousness to include thoughts as well as sensory experiences. But the idea of an important connection between consciousness and thought is an engaging one. Sometimes, for example, it seems hard to accept that there could be a fully satisfying reconstruction of thinking in the terms favored by the physical sciences. This intuition is similar to, and perhaps derives from, the intuition that consciousness somehow defies scientific explanation.

The question whether there is an important connection between consciousness and thought divides into two: Does consciousness require thought? Does thought require consciousness? The intuitive answer to the first question is that access consciousness evidently does require thought, but that phenomenal consciousness does not. (The appeal of this intuitive answer is the source of some objections to the higher-order thought theory of consciousness.) The answer to the second question as it concerns access consciousness is that there is scarcely any distance at all between the notion of thought and the notion of access consciousness. But when we focus on phenomenal consciousness, the answer to the second question is less clear.

John Searle (1990, 586) argues for the connection principle: "The ascription of an unconscious intentional phenomenon to a system implies that the phenomenon is in principle accessible to consciousness." This is to say that, while we can allow for unconscious intentional states, such as unconscious thoughts, these have to be seen as secondary, and as standing in a close relation to conscious intentional states. Searle's argument is naturally interpreted as being directed toward the conclusion that central cases of thinking are at least akin to phenomenally conscious states.

Even if one does not accept Searle's argument for the connection principle, there is a plausible argument for a weaker version of his conclusion. The INTENTIONALITY of human thought involves modes of presentation of objects and properties (see SENSE AND REFERENCE); demonstrative modes of presentation afforded by perceptual experience of objects and their properties constitute particularly clear examples. For example, we think of an object as "that [perceptually presented] cat" or of a property as "that color." Suppose now that it could be argued that some theoretical primacy attaches to these "perceptual demonstrative" modes of presentation (Perry 1979). It might be argued, for example, that in order to be able to think about objects at all, a subject needs to be able to think about objects under perceptual demonstrative modes of presentation. Such an argument would establish a deep connection between intentionality and consciousness.

Finally, there is another way phenomenal consciousness might enter the theory of thought. It might be because a thinker's thoughts are phenomenally conscious states, that they also have the more dispositional properties (such as reportability) mentioned in the definition of access consciousness. This phenomenal consciousness property might also figure in the explanation of a thinker's being able to engage in critical reasoning—evaluating and assessing reasons and reasoning as such (Burge 1996). It is far from clear, however, whether this idea can be worked out in a satisfactory way. Would the idea require a sensational phenomenology for thinking? If it does require that, then it might be natural to suggest that phenomenally conscious thoughts are clothed in the phonological or orthographic forms of natural language sentences (Carruthers 1996).

—Martin Davies

References

Allport, A. (1988). What concept of consciousness? In A. J. Marcel and E. Bisiach, Eds., *Consciousness in Contemporary Science.* Oxford: Oxford University Press, pp. 159–182.

Block, N. (1978). Troubles with functionalism. In C. Wade Savage, Ed., *Minnesota Studies in the Philosophy of Science,* vol. 9. Minneapolis: University of Minnesota Press, pp. 261–325.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18: 227–287.

Braddon-Mitchell, D., and F. Jackson (1996). *Philosophy of Mind and Cognition.* Oxford: Blackwell.

Burge, T. (1996). Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society* 96: 91–116.

Carruthers, P. (1996). *Language, Thought and Consciousness.* Cambridge: Cambridge University Press.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory.* New York: Oxford University Press.

Churchland, P. M. (1985). Reduction, qualia and the direct introspection of brain states. *Journal of Philosophy* 82: 8–28.
Churchland, P. M. (1988). *Matter and Consciousness.* Rev. ed. Cambridge, MA: MIT Press.
Dennett, D. C. (1988). Quining qualia. In A. J. Marcel and E. Bisiach, Eds., *Consciousness in Contemporary Science.* Oxford: Oxford University Press, pp. 42–77.
Dennett, D. C. (1991). *Consciousness Explained.* Boston: Little, Brown.
Flanagan, O. (1992). *Consciousness Reconsidered.* Cambridge, MA: MIT Press.
Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32: 127–136.
Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy* 83: 291–295.
Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly* 64: 354–361.
McGinn, C. (1989). Can we solve the mind-body problem? *Mind* 98: 349–366.
Nagel, T. (1974). What is it like to be a bat? *Philosophical Review* 83: 435–450.
Perry, J. (1979). The problem of the essential indexical. *Noûs* 13: 3–21.
Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology* 47: 44–50.
Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies* 94: 329–359.
Searle, J. R. (1990). Consciousness, explanatory inversion, and cognitive science. *Behavioral and Brain Sciences* 13: 585–596.

Further Readings

Block, N. (1998). How to find the neural correlate of consciousness. In A. O'Hear, Ed., *Contemporary Issues in the Philosophy of Mind.* Royal Institute of Philosophy Supplement 43. Cambridge: Cambridge University Press, pp. 23–34.
Block, N., O. Flanagan, and G. Güzeldere, Eds. (1997). *The Nature of Consciousness: Philosophical Debates.* Cambridge, MA: MIT Press.
Crick, F., and C. Koch. (1995). Are we aware of neural activity in primary visual cortex? *Nature* 375: 121–123.
Davies, M., and G. W. Humphreys, Eds. (1993). *Consciousness: Psychological and Philosophical Essays.* Oxford: Blackwell.
Dennett, D. C., and M. Kinsbourne. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences* 15: 183–247.
Metzinger, T., Ed. (1995). *Conscious Experience.* Paderborn: Schöningh.
Nelkin, N. (1996). *Consciousness and the Origins of Thought.* Cambridge: Cambridge University Press.
Peacocke, C. (1998). Conscious attitudes, attention and self-knowledge. In C. Wright, B. C. Smith, and C. Macdonald, Eds., *Knowing Our Own Minds.* Oxford: Oxford University Press, pp. 63–98.
Rolls, E. T. (1997). Consciousness in neural networks? *Neural Networks* 10: 1227–1240.
Schacter, D. L. (1989). On the relation between memory and consciousness: Dissociable interactions and conscious experience. In H. Roediger and F. Craik, Eds., *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving.* Hillsdale, NJ: Erlbaum.
Shear, J., Ed. (1997). *Explaining Consciousness: The Hard Problem.* Cambridge, MA: MIT Press.
Tye, M. (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind.* Cambridge, MA: MIT Press.

Consciousness, Neurobiology of

After a hiatus of fifty years or more, the physical origins of CONSCIOUSNESS are being once again vigorously debated, in hundreds of books and monographs published in the last decade. What sparse facts can we ascertain about the neurobiological basis of consciousness, and what can we reasonably assume at this point in time?

By and large, neuroscientists have made a number of working assumptions that need to be justified more fully. In particular,

1. There is something to be explained, that is, the subjective content associated with a conscious sensation (what philosophers refer to as QUALIA; see also WHAT-IT'S-LIKE) does exist and has its physical basis in the brain.

2. Consciousness is one of the principal properties of the human brain, a highly evolved system; it must therefore have a useful function to perform. Crick and Koch (1995) assume that the function of visual consciousness is to produce the best current interpretation of the visual scene—in the light of past experiences—and to make it available for a sufficient time to the parts of the brain that contemplate, plan, and execute voluntary motor outputs (including language). This needs to be contrasted with the on-line systems that bypass consciousness but can generate stereotyped behaviors (see below).

3. At least some animal species (i.e., non-human primates such as the macaque monkey) are assumed to possess some aspects of consciousness. Consciousness associated with sensory events is likely to be very similar in humans and monkeys for several reasons. First, trained monkeys behave as humans do under controlled conditions for most sensory tasks (e.g., visual motion discrimination; see MOTION, PERCEPTION OF; Wandell 1995). Second, the gross neuroanatomy of humans and nonhuman primates is the same, once the difference in size has been accounted for. Finally, MAGNETIC RESONANCE IMAGING in humans is confirming the existence of a functional organization very similar to that discovered by single-cell electrophysiology in the monkey (Tootell et al. 1996). As a corollary, it follows that language is not necessary for consciousness to occur (although it greatly enriches human consciousness). In the following, we will mainly concentrate on sensory consciousness, and, in particular, on visual consciousness, because it is experimentally the most accessible and the best understood.

Cognitive and clinical research demonstrates that much complex information processing can occur without involving consciousness, both in normals as well as in patients. Examples of this include BLINDSIGHT (Weiskrantz 1997), priming, and the implicit recognition of complex sequences (Velmans 1991; Berns, Cohen, and Mintun 1997). Milner and Goodale (1995) have made a masterful case for the existence of so-called on-line visual systems that bypass consciousness, and that serve to mediate relative stereotype visual-motor behaviors, such as eye and arm movements as well as posture adjustments, in a very rapid manner. On-line systems work in egocentric coordinate systems and lack both certain types of perceptual ILLUSIONS (e.g. size illusion) and direct access to WORKING MEMORY. Milner and Goodale (1995; see also

Rossetti forthcoming) hypothesize that on-line systems are associated with the dorsal stream of visual information in the CEREBRAL CORTEX, originating in the primary VISUAL CORTEX (V1) and terminating in the posterior parietal cortex (see VISUAL PROCESSING STREAMS). This contrasts well with the function of consciousness alluded to above, namely, to synthesize information from many different sources and use it to plan behavioral patterns over time.

What is the neuronal correlate of consciousness? Most popular has been the belief that consciousness arises as an emergent property of a very large collection of interacting neurons (Popper and Eccles 1981; Libet 1995). An alternative hypothesis is that there are special sets of “consciousness” neurons distributed throughout cortex (and associated systems, such as the THALAMUS and the BASAL GANGLIA) that represent the ultimate neuronal correlate of consciousness (NCC), in the sense that activity of an appropriate subset of them is both necessary and sufficient to give rise to an appropriate conscious experience or percept (Crick and Koch 1995). NCC neurons would, most likely, be characterized by a unique combination of molecular, biophysical, pharmacological, and anatomical traits. It is also possible, of course, that all cortical neurons may be capable of participating in the representation of one percept or another, at one time or another, though not necessarily doing so for all percepts. The secret of consciousness would then consist of all cortical neurons representing that particular percept at that moment (see BINDING BY NEURAL SYNCHRONY).

Where could such NCC neurons be found? Based on clinical evidence that small lesions of the intralaminar nuclei of the thalamus (ILN) cause loss of consciousness and coma and that ILN neurons project widely and reciprocally into the cerebral cortex, ILN neurons have been proposed as the site where consciousness is generated (Bogen 1995; Purpura and Schiff 1997). It is more likely, however, that ILN neurons provide an enabling or arousal signal without which no significant cortical processing can occur. The great specificity associated with the content of our consciousness at any point in time can only be mediated by neurons in the cerebral cortex, its associated specific thalamic nuclei, and the basal ganglia. It is here, among the neurons whose very specific response properties have been extensively characterized by SINGLE-NEURON RECORDING, that we have to look for the NCC.

What, if anything, can we infer about the location of these neurons? In the case of visual consciousness, Crick and Koch (1995) surmised that these neurons must have access to visual information and project to the planning stages of the brain, that is, to premotor and frontal areas (Fuster 1997). Because in the macaque monkey, no neurons in primary visual cortex project to any area anterior to the central sulcus, Crick and Koch (1995) proposed that neurons in V1 do not directly give rise to consciousness (although V1 is necessary for most forms of vision, just as the retina is). Current electrophysiological, psychophysical, and imaging evidence (He, Cavanagh, and Intriligator 1996; Engel, Zhang, and Wandell 1997) supports the hypothesis that the NCC is not to be found among V1 neurons.

A promising experimental approach to locate the NCC has been the use of bistable percepts, that is, pairs of percepts, alternating in time, that arise from a constant visual stimulus as in a Necker cube (Crick and Koch 1992). In one such case, a small image, say of a horizontal grating, is presented to the left eye and another image, say of a vertical grating, is presented to the corresponding location in the right eye. In spite of the constant retinal stimulus, observers “see” the horizontal grating alternate every few seconds with the vertical one, a phenomenon known as “binocular rivalry” (Blake 1989). The brain does not allow for the simultaneous perception of both images.

It is possible, though difficult, to train a macaque monkey to report whether it is currently seeing the left or the right image. The distribution of the switching times and the way in which changing the contrast in one eye affects these times leaves little doubt that monkeys and humans experience the same basic phenomenon (Myerson, Miezin, and Allman 1981). In a series of elegant experiments, Logothetis and colleagues (Logothetis and Schall 1989; Leopold and Logothetis 1996; Sheinberg and Logothetis 1997) recorded from a variety of monkey cortical areas during this task. In early visual cortex, only a small fraction of cells modulated their response as a function of the percept of the monkey, while 20 to 30 percent of neurons in MT and V4 cells did. The majority of cells increased their firing rate in response to one or the other retinal stimulus with no regard to what the animal perceived at the time. In contrast, in a high-level cortical area, such as the inferior temporal cortex (IT), almost all neurons responded only to the perceptual dominant stimulus (in other words, a “face” cell only fired when the animal indicated by its performance that it saw the face and not the sunburst pattern in the other eye). This makes it likely that the NCC is located among—or beyond—IT neurons.

Finding the NCC would only be the first, albeit critical, step in understanding consciousness. We also need to know where these cells project to, their postsynaptic action, and what happens to them in various diseases known to affect consciousness, such as schizophrenia or AUTISM, and so on. And, of course, a final theory of consciousness would have to explain the central mystery—why a physical system with a particular architecture gives rise to feelings and qualia. (Chalmers 1996).

See also ATTENTION; ATTENTION IN THE ANIMAL BRAIN; ATTENTION AND THE HUMAN BRAIN; SENSATIONS

—*Christof Koch and Francis Crick*

References

Berns, G. S., J. D. Cohen, and M. A. Mintun. (1997). Brain regions responsive to novelty in the absence of awareness. *Science* 276: 1272–1275.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychol. Rev.* 96: 145–167.

Bogen, J. E. (1995). On the neurophysiology of consciousness: 1. An overview. *Consciousness and Cognition* 4: 52–62.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory.* Oxford: Oxford University Press.

Crick, F., and C. Koch. (1992). The problem of consciousness. *Scientific American* 267(3): 153–159.

Crick, F., and C. Koch. (1995). Are we aware of neural activity in primary visual cortex? *Nature* 375: 121–123.

Engel, S., X. Zhang, and B. Wandell (1997). Colour tuning in human visual cortex measured with functional magnetic resonance imaging. *Nature* 388: 68–71.

Fuster, J. M. (1997). *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe.* 3rd ed. Philadelphia: Lippincott-Raven.

He, S., P. Cavanagh, and J. Intriligator. (1996). Attentional resolution and the locus of visual awareness. *Nature* 383: 334–337.

Leopold, D. A., and N. K. Logothetis. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379: 549–553.

Libet, B. (1995). *Neurophysiology of Consciousness: Selected Papers and New Essays.* Boston: Birkhäuser.

Logothetis, N., and J. Schall. (1989). Neuronal correlates of subjective visual perception. *Science* 245: 761–763.

Milner, D., and M. Goodale. (1995). *The Visual Brain in Action.* Oxford: Oxford University Press.

Myerson, J., F. Miezin, and J. Allman. (1981). Binocular rivalry in macaque monkeys and humans: A comparative study in perception. *Behav. Anal. Lett.* 1: 149–156.

Popper, K. R., and J. C. Eccles. (1981). *The Self and Its Brain.* Berlin: Springer.

Purpura, K. P., and N. D. Schiff (1997). The thalamic intralaminar nuclei: a role in visual awareness. *Neuroscientist* 3: 8–15.

Rossetti, Y. (Forthcoming). Implicit perception in action: Short-lived motor representations of space evidenced by brain-damaged and healthy subjects. In P. G. Grossenbacher, Ed., *Finding Consciousness in the Brain.* Philadelphia: Benjamins.

Sheinberg, D. L., and N. K. Logothetis. (1997). The role of temporal cortical areas in perceptual organization. *Proc. Natl. Acad. Sci. U.S.A.196* 94: 3408–3413.

Tootell, R. B. H., A. M. Dale, M. I. Sereno, and R. Malach. (1996). New images from human visual cortex. *Trends Neurosci.* 19: 481–489.

Velmans, M. (1991). Is human information processing conscious? *Behavioral Brain Sci.* 14: 651–726.

Wandell, B. A. (1995). *Foundations of Vision.* Sunderland, MA: Sinauer.

Weiskrantz, L. (1997). *Consciousness Lost and Found.* Oxford: Oxford University Press.

Further Readings

Crick, F., and C. Koch. (1998). Consciousness and neuroscience. *Cerebral Cortex* 8: 97–107.

Jackendoff, R. (1987). *Consciousness and the Computational Mind.* Cambridge, MA: MIT Press.

Zeki, S. (1993). *Vision of the Brain.* Oxford: Blackwell.

Consensus Theory

See CULTURAL CONSENSUS THEORY

Constraint Satisfaction

A *constraint satisfaction problem* (CSP) is defined over a *constraint network,* which consists of a finite set of *variables,* each associated with a *domain* of values, and a set of *constraints*. A *solution* is an assignment of a value to each variable from its domain such that all the constraints are satisfied. Typical constraint satisfaction problems are to determine whether a solution exists, to find one or all solutions, and to find an optimal solution relative to a given cost function. A well-known example of a constraint satisfaction problem is *k*-colorability, where the task is to color, if possible, a given graph with *k* colors only, such that any two adjacent nodes have different colors. A constraint satisfaction formulation of this problem associates the nodes of the graph with variables, the possible colors are their domains, and the inequality constraints between adjacent nodes are the constraints of the problem. Each constraint of a CSP may be expressed as a relation, defined on some subset of variables, denoting legal combinations of their values. Constraints can also be described by mathematical expressions or by computable procedures. Another typical constraint satisfaction problem is *SATisfiability,* the task of finding the truth assignment to propositional variables such that a given set of clauses is satisfied. For example, given the two clauses $(A \vee B \vee \neg C)$, $(\neg A \vee D)$, the assignment of *false* to *A*, *true* to *B*, *false* to *C*, and *false* to *D*, is a satisfying truth value assignment.

The structure of a constraint network is depicted by a constraint graph whose nodes represent the variables and in which any two nodes are connected if the corresponding variables participate in the same constraint. In the *k*-colorability formulation, the graph to be colored is the constraint graph. In our SAT example the constraint graph has *A* connected to *D*, and *A*, *B*, and *C* are connected to each other.

Constraint networks have proven successful in modeling mundane cognitive tasks such as vision, language comprehension, default reasoning, and abduction, as well as in applications such as scheduling, design, diagnosis, and temporal and spatial reasoning. In general, constraint satisfaction tasks are computationally intractable ("NP-hard"; see COMPUTATIONAL COMPLEXITY).

ALGORITHMS for processing constraints can be classified into two interacting categories: (1) search and (2) consistency inference. Search algorithms traverse the space of partial instantiations, while consistency inference algorithms reason through equivalent problems. Search algorithms are either systematic and complete or stochastic and incomplete. Likewise, consistency inference algorithms have either complete solutions (e.g., variable-elimination algorithms) or incomplete solutions (i.e., local consistency algorithms).

Local consistency algorithms, also called "consistency-enforcing" or "constraint propagation" algorithms (Montanari 1974; Mackworth 1977; Freuder 1982), are polynomial algorithms that transform a given constraint network into an equivalent, yet more explicit network by deducing new constraints to be added onto the network. Intuitively, a consistency-enforcing algorithm will make any partial solution of a small subnetwork extensible to some surrounding network. For example, the most basic consistency algorithm, called an "arc consistency" algorithm, ensures that any legal value in the domain of a single variable has a legal match in the domain of any other selected variable. A "path consistency" algorithm ensures that any consistent solution to a two-variable subnetwork is extensible to any third variable, and, in general, *i-consistency* algorithms guarantee that any locally consistent instantiation of $i - 1$ variables is extensible to any *i*th variable. Enforcing *i*-consistency is time and space