The Puzzle of Conscious Experience

David J. Chalmers

Department of Philosophy
University of California
Santa Cruz, CA 63130

chalmers@paradox.ucsc.edu

[*Scientific American*, December 1995 pp. 62-68. N.B. As always at *Scientific American*, this isn't just me talking. For a more careful treatment of this material, see my "Facing Up to the Problem of Consciousness".]

Conscious experience is at once the most familiar thing in the world and the most mysterious. There is nothing we know about more directly than consciousness, but it is extraordinarily hard to reconcile it with everything else we know. Why does it exist? What does it do? How could it possibly arise from neural processes in the brain? These questions are among the most intriguing in all of science.

From an objective viewpoint, the brain is relatively comprehensible. When you look at this page, there is a whir of processing: photons strike your retina, electrical signals are passed up your optic nerve and between different areas of your brain, and eventually you might respond with a smile, a perplexed frown or a remark. But there is also a subjective aspect. When you look at the page, you are conscious of it, directly experiencing the images and words as part of your private, mental life. You have vivid impressions of colored flowers and vibrant sky. At the same time, you may be feeling some emotions and forming some thoughts. Together such experiences make up consciousness: the subjective, inner life of the mind.

For many years, consciousness was shunned by researchers studying the brain and the mind. The prevailing view was that science, which depends on objectivity, could not accommodate something as subjective as consciousness. The behaviorist movement in psychology, dominant earlier in this century concentrated on external behavior and disallowed any talk of internal mental processes. Later, the rise of cognitive science focused attention on processes inside the head. Still, consciousness remained off-limits, fit only for late-night discussion over drinks.

Over the past several years, however, an increasing number of neuroscientists, psychologists and philosophers have been rejecting the idea that consciousness cannot be studied and are attempting to delve into its secrets. As might be expected of a field so new, there is a tangle of diverse and conflicting theories, often using basic concepts in incompatible ways. To help unsnarl the tangle, philosophical reasoning is vital.

The myriad views within the field range from reductionist theories, according to which consciousness can be explained by the standard methods of neuroscience and psychology, to the position of the so-called mysterians, who say we will never understand consciousness at all. I believe that on close analysis both of these views can be seen to be mistaken and that the truth lies somewhere in the middle. Against reductionism I will argue that the tools of neuroscience cannot provide a full account of conscious experience, although they have much to offer. Against mysterianism I will hold that consciousness might be explained by a new kind of theory. The full details of such a theory are still out of reach, but careful reasoning and some educated inferences can reveal something of its general nature. For example, it will probably involve new fundamental laws, and the concept of information may play a central role. These faint glimmerings suggest that a theory of consciousness may have startling consequences for our view of the universe and of ourselves.

The Hard Problem

Researchers use the word "consciousness" in many different ways. To clarify the issues, we first have to separate the problems that are often clustered together under the name. For this purpose, I find it useful to distinguish between the "easy problems" and the "hard problem" of consciousness. The easy problems are by no means trivial - they are actually as challenging as most in psychology and biology - but it is with the hard problem that the central mystery lies.

The easy problems of consciousness include the following: How can a human subject discriminate sensory stimuli and react to them appropriately? How does the brain integrate information from many different sources and use this information to control behavior? How is it that subjects can verbalize their internal states? Although all these questions are associated with consciousness, they all concern the objective mechanisms of the cognitive system. Consequently, we have every reason to expect that continued work in cognitive psychology and neuroscience will answer them.

The hard problem, in contrast, is the question of how physical processes in the brain give rise to subjective experience. This puzzle involves the inner aspect of thought and perception: the way things feel for the subject. When we see, for example, we experience visual sensations, such as that of vivid blue. Or think of the ineffable sound of a distant oboe, the agony of an intense pain, the sparkle of happiness or the meditative quality of a moment lost in thought. All are part of what I am calling consciousness. It is these phenomena that pose the real mystery of the mind.


ISOLATED NEUROSCIENTIST in a black-and-white room knows everything about how the brain processes colors but does not know what it is like to see them. This scenario suggests that knowledge of the brain does not yield complete knowledge of conscious experience.

To illustrate the distinction, consider a thought experiment devised by the Australian philosopher Frank Jackson. Suppose that Mary, a neuroscientist in the 23rd century, is the world's leading expert on the brain processes responsible for color vision. But Mary has lived her whole life in a black-and-white room and has never seen any other colors. She knows everything there is to know about physical processes in the brain - its biology, structure and function. This understanding enables her to grasp everything there is to know about the easy problems: how the brain discriminates stimuli, integrates information and produces verbal reports. From her knowledge of color vision, she knows the way color names correspond with wavelengths on the light spectrum. But there is still something crucial about color vision that Man does not know: what it is like to experience a color such as red. It follows that there are facts about conscious experience that cannot be deduced from physical facts about the functioning of the brain.

Indeed, nobody knows why these physical processes are accompanied by conscious experience at all. Why is it that when our brains process light of a certain wavelength, we have an experience of deep purple? Why do we have any experience at all? Could not an unconscious automaton have performed the same tasks just as well? These are questions that we would like a theory of consciousness to answer.

I am not denying that consciousness arises from the brain. We know, for example, that the subjective experience of vision is closely linked to processes in the visual cortex. It is the link itself that perplexes, however. Remarkably, subjective experience seems to emerge from a physical process. But we have no idea how or why this is.

Is Neuroscience Enough?

Given the flurry of recent work on Consciousness in neuroscience and psychology, one might think this mystery is starting to be cleared up. On closer examination, however, it turns out that almost all the current work addresses only the easy problems of consciousness. The confidence of the reductionist view comes from the progress on the easy problems, but none of this makes any difference where the hard problem is concerned.

Consider the hypothesis put forward by neurobiologists Francis Crick of the Salk Institute for Biological Studies in San Diego and Christof Koch of the California Institute of Technology. They suggest that consciousness may arise from certain oscillations in the cerebral cortex, which become synchronized as neurons fire 40 times per second. Crick and Koch believe the phenomenon might explain how different attributes of a single perceived object (its color and shape, for example), which are processed in different parts of the brain, are merged into a coherent whole. In this theory, two pieces of information become bound together precisely when they are represented by synchronized neural firings.

The hypothesis could conceivably elucidate one of the easy problems about how information is integrated in the brain. But why should synchronized oscillations give rise to a visual experience, no matter how much integration is taking place? This question involves the hard problem, about which the theory has nothing to offer. Indeed, Crick and Koch are agnostic about whether the hard problem can be solved by science at all.

The same kind of critique could be applied to almost all the recent work on consciousness. In his 1991 book *Consciousness Explained*, philosopher Daniel C. Dennett laid out a sophisticated theory of how numerous independent processes in the brain combine to produce a coherent response to a perceived event. The theory might do much to explain how we produce verbal reports on our internal states, but it tells us very little about why there should be a subjective experience behind these reports. Like other reductionist theories, Dennett's is a theory of the easy problems.

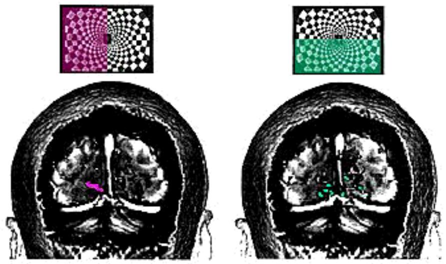
The critical common trait among these easy problems is that they all concern how a cognitive or behavioral function is performed. All are ultimately questions about how the brain carries out some task-how it discriminates stimuli, integrates information, produces reports and so on. Once neurobiology specifies appropriate neural mechanisms, showing how the functions are performed, the easy problems are solved. The hard problem of consciousness, in contrast, goes beyond problems about how functions are performed. Even if every behavioral and cognitive function related to consciousness were explained, there would still remain a further mystery: Why is the performance of these functions accompanied by conscious experience? It is this additional conundrum that makes the hard problem hard.

The Explanatory Gap

Some have suggested that to solve the hard problem, we need to bring in new tools of physical explanation: nonlinear dynamics, say, or new discoveries in neuroscience, or quantum mechanics. But these ideas suffer from exactly the same difficulty. Consider a proposal from Stuart R. Hameroff of the University of Arizona and Roger Penrose of the University of Oxford. They hold that consciousness arises from quantum-physical processes taking place in microtubules, which are protein structures inside neurons. It is possible (if not likely) that such a hypothesis will lead to an explanation of how the brain makes decisions or even how it proves mathematical theorems, as Hameroff and Penrose suggest. But even if it does, the theory is silent about how these processes might give rise to conscious experience. Indeed, the same problem arises with any theory of consciousness based only on physical processing.

The trouble is that physical theories are best suited to explaining why systems have a certain physical structure and how they perform various functions. Most problems in science have this form; to explain life, for example, we need to describe how a physical system can reproduce, adapt and metabolize. But consciousness is a different sort of problem entirely, as it goes beyond the explanation of structure and function.

Of course, neuroscience is not irrelevant to the study of consciousness. For one, it may be able to reveal the nature of the neural correlate of consciousness - the brain processes most directly associated with conscious experience. It may even give a detailed correspondence between specific processes in the brain and related components of experience. But until we know why these processes give rise to conscious experience at all, we will not have crossed what philosopher Joseph Levine has called the explanatory gap between physical processes and consciousness. Making that leap will demand a new kind of theory.



BLOOD FLOW variations in the visual cortex demonstrate how a subject's brain responds to a pattern being viewed. The colors in this image show the cortical activity corresponding to the subject's view of either half of the pattern. This activity is a neural correlate of visual experience, but the relationship between the two remains mysterious.

A True Theory of Everything

In searching for an alternative, a key observation is that not all entities in science are explained in terms of more basic entities. In physics, for example, space-time, mass and charge (among other things) are regarded as fundamental features of the world, as they are not reducible to anything simpler. Despite this irreducibility, detailed and useful theories relate these entities to one another in terms of fundamental laws. Together these features and laws explain a great variety of complex and subtle phenomena.

It is widely believed that physics provides a complete catalogue of the universe's fundamental features and laws. As physicist Steven Weinberg puts it in his 1992 book *Dreams of a Final Theory*, the goal of physics is a "theory of everything" from which all there is to know about the universe can be derived. But Weinberg concedes that there is a problem with consciousness. Despite the power of physical theory, the existence of consciousness does not seem to be derivable from physical laws. He defends physics by arguing that it might eventually explain what he calls the objective correlates of consciousness (that is, the neural correlates), but of course to do this is not to explain consciousness itself. If the existence of consciousness cannot be derived from physical laws, a theory of physics is not a true theory of everything. So a final theory must contain an additional fundamental component.

Toward this end, I propose that conscious experience be considered a fundamental feature, irreducible to anything more basic. The idea may seem strange at first, but consistency seems to demand it. In the 19th century it turned out that electromagnetic phenomena could not be explained in terms of previously known principles. As a consequence, scientists introduced electromagnetic charge as a new fundamental entity and studied the associated fundamental laws. Similar reasoning should apply to consciousness. If existing fundamental theories cannot encompass it, then something new is required.

Where there is a fundamental property, there are fundamental laws. In this case, the laws must relate experience to elements of physical theory. These laws will almost certainly not interfere with those of the physical world; it seems that the latter form a closed system in their own right. Rather the laws will serve as a bridge, specifying how experience depends on underlying physical processes. It is this bridge that will cross the explanatory gap.

Thus, a complete theory will have two components: physical laws, telling us about the behavior of physical systems from the infinitesimal to the cosmological, and what we might call psychophysical laws, telling us how some of those systems are associated with conscious experience. These two components will constitute a true theory of everything.

Searching for a Theory

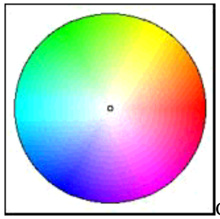
Supposing for the moment that they exist, how might we uncover such psychophysical laws? The greatest hindrance in this pursuit will be a lack of data. As I have described it, consciousness is subjective, so there is no direct way to monitor it in others. But this difficulty is an obstacle, not a dead end. For a start, each one of us has access to our own experiences, a rich trove that can be used to formulate theories. We can also plausibly rely on indirect information, such as subjects' descriptions of their experiences. Philosophical arguments and thought experiments also have a role to play. Such methods have limitations, but they give us more than enough to get started.

These theories will not be conclusively testable, so they will inevitably be more speculative than those of more conventional scientific disciplines. Nevertheless, there is no reason they should not be strongly constrained to account accurately for our own first-person experiences, as well as the evidence from subjects' reports. If we find a theory that fits the data better than any other theory of equal simplicity, we will have good reason to accept it. Right now we do not have even a single theory that fits the data, so worries about testability are premature.

We might start by looking for high-level bridging laws, connecting physical processes to experience at an everyday level. The basic contour of such a law might be gleaned from the observation that when we are conscious of something, we are generally able to act on it and speak about it - which are objective, physical functions. Conversely, when some information is directly available for action and speech, it is generally conscious. Thus, consciousness correlates well with what we might call "awareness": the process by which information in the brain is made globally available to motor processes such as speech and bodily action.

The notion may seem trivial. But as defined here, awareness is objective and physical, whereas consciousness is not. Some refinements to the definition of awareness are needed, in order to extend the concept to animals and infants, which cannot speak. But at least in familiar cases, it is possible to see the rough outlines of a psychophysical law: where there is awareness, there is consciousness, and vice versa.

To take this line of reasoning a step further, consider the structure present in the conscious experience. The experience of a field of vision, for example, is a constantly changing mosaic of colors, shapes and patterns and as such has a detailed geometric structure. The fact that we can describe this structure, reach out in the direction of many of its components and perform other actions that depend on it suggests that the structure corresponds directly to that of the information made available in the brain through the neural processes of awareness.



COLOR WHEEL arranges hues so that ones experienced as similar are closest. Nearby colors also correspond to perceptual representations in the brain.

Similarly, our experiences of color have an intrinsic three-dimensional structure that is mirrored in the structure of information processes in the brain's visual cortex. This structure is illustrated in the color wheels and charts used by artists. Colors are arranged in a systematic pattern - red to green on one axis, blue to yellow on another, and black to white on a third. Colors that are close to one another on a color wheel are experienced as similar. It is extremely likely that they also correspond to similar perceptual representations in the brain, as part of a system of complex three-dimensional coding among neurons that is not yet fully understood. We can recast the underlying concept as a principle of structural coherence: the structure of conscious experience is mirrored by the structure of information in awareness, and vice versa.

Another candidate for a psychophysical law is a principle of organizational invariance. It holds that physical systems with the same abstract organization will give rise to the same kind of conscious experience, no matter what they are made of. For example, if the precise interactions between our neurons could be duplicated with silicon chips, the same conscious experience would arise. The idea is somewhat controversial, but I believe it is strongly supported by thought experiments describing the gradual replacement of neurons by silicon chips. The remarkable implication is that consciousness might someday be achieved in machines.

Information: Physical and Experiential

The ultimate goal of a theory of consciousness is a simple and elegant set of fundamental laws, analogous to the fundamental laws of physics. The principles described above are unlikely to be fundamental, however. Rather they seem to be high-level psychophysical laws, analogous to macroscopic principles in physics such as those of thermodynamics or kinematics. What might the underlying fundamental laws be? No one knows, but I don't mind speculating.

I suggest that the primary psychophysical laws may centrally involve the concept of information. The abstract notion of information, as put forward in the 1940s by Claude E. Shannon of the Massachusetts Institute of Technology, is that of a set of separate states with a basic structure of similarities and differences between them. We can think of a 10-bit binary code as an information state, for example. Such information states can be embodied in the physical world. This happens whenever they correspond to physical states (voltages, say), the differences between which can be transmitted along some pathway, such as a telephone line.

We can also find information embodied in conscious experience. The pattern of color patches in a visual field, for example, can be seen as analogous to that of the pixels covering a display screen. Intriguingly, it turns out that we find the same information states embedded in conscious experience and in underlying physical processes in the brain. The three-dimensional encoding of color spaces, for example, suggests that the information state in a color experience corresponds directly to an information state in the brain. We might even regard the two states as distinct aspects of a single information state, which is simultaneously embodied in both physical processing and conscious experience.

A natural hypothesis ensues. Perhaps information, or at least some information, has two basic aspects: a physical one and an experiential one. This hypothesis has the status of a fundamental principle that might underlie the relation between physical processes and experience. Wherever we find conscious experience, it exists as one aspect of an information state, the other aspect of which is embedded in a physical process in the brain. This proposal needs to be fleshed out to make a satisfying theory. But it fits nicely with the principles mentioned earlier - systems with the same organization will embody the same information, for example - and it could explain numerous features of our conscious experience.

The idea is at least compatible with several others, such as physicist John A. Wheeler's suggestion that information is fundamental to the physics of the universe. The laws of physics might ultimately be cast in informational terms, in which case we would have a satisfying congruence between the constructs in both physical and psychophysical laws. It may even be that a theory of physics and a theory of consciousness could eventually be consolidated into a single grander theory of information.

A potential problem is posed by the ubiquity of information. Even a thermostat embodies some information, for example, but is it conscious? There are at least two possible responses. First, we could constrain the fundamental laws so that only some information has an experiential aspect, perhaps depending on how it is physically processed. Second, we might bite the bullet and allow that all information has an experiential aspect - where there is complex information processing, there is complex experience, and where there is simple information processing, there is simple experience. If this is so, then even a thermostat might have experiences, although they would be much simpler than even a basic color experience, and there would certainly be no accompanying emotions or thoughts. This seems odd at first, but if experience is truly fundamental, we might expect it to be widespread. In any case, the choice between these alternatives should depend on which can be integrated into the most powerful theory.

Of course, such ideas may be all wrong. On the other hand, they might evolve into a more powerful proposal that predicts the precise structure of our conscious experience from physical processes in our brains. If this project succeeds, we will have good reason to accept the theory. If it fails, other avenues will be pursued, and alternative fundamental theories may be developed. In this way, we may one day resolve the greatest mystery of the mind.

Dancing Qualia in a Synthetic Brain [Box]

Whether consciousness could arise in a complex, synthetic system is a question many people find intrinsically fascinating. Although it may be decades or even centuries before such a system is built, a simple thought experiment offers strong evidence that an artificial brain, if organized appropriately, would indeed have precisely the same kind of conscious experiences as a human being.

Consider a silicon-based system in which the chips are organized and function in the same way as the neurons in your brain. That is, each chip in the silicon system does exactly what its natural analogue does and is interconnected to surrounding elements in precisely the same way. Thus, the behavior exhibited by the artificial system will be exactly the same as yours. The crucial question is: Will it be conscious in the same way that you are?

Let us assume, for the purpose of argument, that it would not be. (Here we use a reasoning technique known as reductio ad absurdum, in which the opposite hypothesis is assumed and then shown to lead to an untenable conclusion.) That is, it either has different experiences - an experience of blue, say, when you are seeing red - or no experience at all. We will consider the first case; the reasoning proceeds similarly in both cases.

Because chips and neurons have the same function, they are interchangeable, with the proper interfacing. Chips therefore can replace neurons, producing a continuum of cases in which a successively larger proportion of neurons are replaced by chips. Along this continuum, the conscious experience of the system will also change. For example, we might replace all the neurons in your visual cortex with an identically organized version made of silicon. The resulting brain, with an artificial visual cortex, will have a different conscious experience from the original: where you had previously seen red, you may now experience purple (or perhaps a faded pink, in the case where the wholly silicon system has no experience at all).

Both visual cortices are then attached to your brain, through a two-position switch. With the switch in one mode, you use the natural visual cortex; in the other, the artificial cor tex is activated. When the switch is flipped, your experience changes from red to purple, or vice versa. When the switch is flipped repeatedly, your experiences 'dance' between the two different conscious states (red and purple), known as qualia.

Because your brain's organization has not changed, however, there can be no behavioral change when the switch is thrown. Therefore, when asked about what you are seeing, you will say that nothing has changed. You will hold that you are seeing red and have seen nothing but red, even though the two colors are dancing before your eyes. This conclusion is so unreasonable that it is best taken as a reductio ad absurdum of the original assumption-that an artificial system with identical organization and functioning has a different conscious experience from that of a neural brain. Retraction of the assumption establishes the opposite: that systems with the same organization have the same conscious experience. -D.J.C.

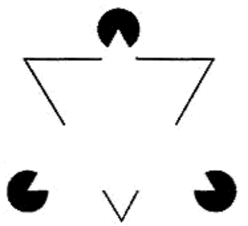
Why Neuroscience May Be Able to Explain Consciousness [Box]

by Francis Crick and Christof Koch

We believe that at the moment the best approach to the problem of explaining consciousness Is to concentrate on finding what is known as the neural correlates of consciousness the processes in the brain that are most directly responsible for consciousness. By locating the neurons In the cerebral cortex that correlate best with consciousness, and figuring out how they link to neurons elsewhere in the brain, we may come across key insights into what David J. Chalmers calls the hard problem: a full accounting of the manner in which subjective experience arises from these cerebral processes.

We commend Chalmers for boldly recognizing and focusing on the hard problem at this early stage, although we are not as enthusiastic about some of his thought experiments. As we see it, the hard problem can be broken down into several questions: Why do we experience anything at all? What leads to a particular conscious experience (such as the blueness of blue)? Why are some aspects of subjective experience impossible to convey to other people (in other words, why are they private)? We believe we have an answer to the last problem and a suggestion about the first two, revolving around a phenomenon known as explicit neuronal representation.

What does 'explicit' mean in this context? Perhaps the best way to define it is with an example. In response to the image of a face, say, ganglion cells fire all over the retina, much like the pixels on a television screen, to generate an implicit representation of the face. At the same time, they can also respond to a great many other features in the image, such as shadows, lines, uneven lighting and so on. In contrast, some neurons high in the hierarchy of the visual cortex respond mainly to the face or even to the face viewed at a particular angle. Such neurons help the brain represent the face in an explicit manner. Their loss resulting from a stroke or some other injury, leads to prosopagnosia, an individual's inability to recognize familiar faces, consciously-even his or her own, although the person can still identify a face as a face. Similarly, damage to other parts of the visual cortex can cause someone to lose the ability experience color, while still seeing In shades of black and white, even though there is no defect in the color receptors in the eye.



KANIZSA TRIANGLE stimulates neurons that code explicitly for such illusory contours.

At each stage, visual information is re-encoded, typically in a semi-hierarchical manner. Retinal ganglion cells respond to areas of light. Neurons in the primary visual cortex are most adept at responding to lines to edges; neurons higher up might prefer a moving contour. Still higher are those that respond to faces and other familiar objects. On top are those that project to pre-motor and motor structures in the brain, where they fire the neurons that initiate such actions as speaking or avoiding an oncoming automobile.

Chalmers believes, as we do, that the subjective aspect of an experience must relate closely to the firing of the neurons corresponding to those aspects (the neural correlates). He describes a well-known thought experiment, constructed around a hypothetical neuroscientist, Mary, who specializes in color perception but has never seen - a color. We believe the reason Mary does not know what it is like to see a color, however, is that she has never had an explicit neural representation of a color in her brain, only of the words and ideas associated with colors.

In order to describe a subjective visual experience, the information has to be transmitted to the motor output stage of the brain, where it becomes available for verbalization or other actions. This transmission always involves re-encoding the information, so that the explicit information expressed by the motor neurons is related, but not identical, to the explicit information expressed by the neurons associated with color experience, at some level in the visual hierarchy.

It is not possible, then, to convey with words and ideas the exact nature of a subjective experience. It is possible, however, to convey a difference between subjective experiences-to distinguish between red and orange, for example. This is possible because a difference in high-level visual cortical area will still be associated with a difference in the motor stages. The implication is that we can never explain to other people the nature of any conscious experience, only its relation to other ones.

The other two questions, concerning why we have conscious experiences and what leads to specific ones, appear more difficult. Chalmers proposes that they require the introduction of 'experience' as a fundamental new feature of the world, relating to the ability of an organism to process information. But which types of neuronal information produce consciousness? And what makes a certain type of information correspond to the blueness of blue, rather than the greenness of green? Such problems seem as difficult as any in the study of consciousness.

We prefer an alternative approach, involving the concept of 'meaning.' In what sense can neurons that explicitly code for a face be said to convey the meaning of a face to the rest of the brain? Such a property must relate to the cell's projective field a pattern of synaptic connections to neurons that code explicitly for related concepts. Ultimately, these connections extend to the motor output. For example, neurons responding to a certain face might be connected to ones expressing the name of the person whose face it is and to others for her voice, memories involving her and so on. Such associations among neurons must be behaviorally useful, in other words, consistent with feedback from the body and the external world.

Meaning derives from the linkages among these representations with others spread throughout the cortical system in a vast associational network, similar to a dictionary or a relational database. The more diverse these connections, the richer the meaning. If, as in our previous example of prosopagnosia, the synaptic output of such face neurons were blocked, the cells would still respond to the person's face, but there would be no associated meaning and, therefore, much less experience. A face would be seen but not recognized as such.

Of course, groups of neurons can take on new functions, allowing brains to learn new categories (including faces) and associate new categories with existing ones. Certain primitive associations, such as pain, are to some extent inborn but subsequently refined in life.

Information may indeed be the key concept, as Chalmers suspects. Greater certainty will require consideration of highly parallel streams of information, linked-as are neurons-in complex networks. It would be useful to try to determine what features a neural network (or some other such computational embodiment) must have to generate meaning. It is possible that such exercises will suggest the neural basis of meaning. The hard problem of consciousness may then appear in an entirely new light. It might even disappear.

The Author

David J. Chalmers studied mathematics at Adelaide University and as a Rhodes Scholar at the University of Oxford, but a fascination with consciousness led him into philosophy and cognitive science. He has a Ph.D. in these fields from Indiana University and is currently in the department of philosophy at the University of California Santa Cruz. Chalmers has published numerous articles on artificial intelligence and the philosophy of mind. His book *The Conscious Mind*, which elaborates many of the ideas in this article, is forthcoming from Oxford University Press.

Further Reading

Absent Qualia, Fading Qualia, Dancing Qualia. David J. Chalmers in *Conscious Experience*. Edited by Thomas Metzinger. Ferdinand Schoningh,1995.

Explaining Consciousness: The Hard Problem. Special issue of Journal of Consciousness Studies, Vol. 2, No. 3; Autumn 1995.

The Nature of Consciousness: Philosophical and Scientific Debates. Edited by Ned Block, Owen Flanagan and Güven Güzeldere. MIT Press (in press).