

Interview with Chalmers

David Chalmers is currently Professor of Philosophy at the University of California at Santa Cruz, and as of January 1999, he will be Professor of Philosophy and Associate Director of the Center for Consciousness Studies at the University of Arizona. He is the author of the well-known book, [*The Conscious Mind*](#) (1996), and is the subject of a collection of critical essays on this book: [*Explaining Consciousness: The Hard Problem*](#) (edited by Jonathan Shear, 1997). He also manages a [web site](#) which includes the most exhaustive [bibliography on philosophy of mind](#) ever assembled, a directory of various online papers pertaining to consciousness, his own papers, and his responses to his critics.

Andrew Chrucky is an independent scholar with a Ph.D. in philosophy. He manages two web sites: [Problems from Wilfrid Sellars](#), which contains exhaustive bibliographies of work by and related to Sellars -- including the texts of many papers and portions of books. His other site, [Digital Text International](#), features the [Meta-Encyclopedia of Philosophy](#), which links six Encyclopedias and Dictionaries under one set of entries.

Chrucky: In the last few years, we have witnessed all sorts of books and articles written on the subject of consciousness, and from among these, I think your book *The Conscious Mind* has provoked the most controversy in a very short time -- to the extent that there already has been published a book of critical essays on your work, with a response by you. Could you explain this proliferation of literature on consciousness, and why this focus on your particular book?

Chalmers: That's a hard question to answer. I think that consciousness has always been the most important topic in the philosophy of mind, and one of the most important topics in cognitive science as a whole, but it had been surprisingly neglected in recent years. But people in the field have always been conscious of that neglect in the back of their minds, so when a few people like Francis Crick and Daniel Dennett started drawing attention to the topic a few years ago, it didn't take much to start a flood. Within psychology and neuroscience, some new and rigorous experimental paradigms for studying consciousness have helped it begin to overcome the stigma that has been attached to the topic for most of this century. And in philosophy, where this stigma has never really been present, I think people have begun to realize that there's no hope of a serious and satisfying philosophy of mind without addressing the problem of consciousness head-on.

As for my own work, much of it is due to being in the right place at the right time. Despite the rush of work on consciousness in the first part of this decade, I think a lot of people had a nagging feeling that the central issues weren't being directly addressed, and I may have helped articulate this feeling with respect to the sciences and the general public. Of course this was already a familiar theme in philosophy, but for some reason that voice wasn't being heard by the wider audience. It probably helps that my background is in the sciences and I can speak the scientists' language.

Within philosophy, I think the reasons are a bit different. Here, the broader issues are already familiar, and discussion has focused at a more sophisticated and detailed level. Within the philosophy of mind, the problem of consciousness is no big news. All the same, materialism -- the view that the mental is nothing more than the physical -- has become a received wisdom in recent years, even though there has been a sort of unease about consciousness in the background. People have managed to avert their eyes and hope for the best. There have been a number of people putting forward anti-materialist views, but these have been a little piecemeal, often in the form of bite-sized articles rather than detailed research programs. In my book I try to give some really systematic arguments, drawing on a background framework in metaphysics and the philosophy of language, for why materialism and the existence of consciousness can't be reconciled, and then I try to outline in some detail where we should go from there. I suppose it is a combination of the controversial nature of the thesis and the details of the framework that

interest people.

Chrucky: I understand that your background is in science. What made you concentrate on such a topic as consciousness from a philosophical perspective?

Chalmers: Even when I was studying mathematics, physics, and computer science, it always seemed that the problem of consciousness was about the most interesting problem out there for science to come to grips with. I found myself thinking about it more and more, until I decided that I wanted to work on it full time. The sort of work I wanted to do couldn't really be done in psychology or neuroscience, where experimental details are what really count and the rest is incidental. So I threw myself into learning philosophy. At first I was very impatient with the kind of nit-picky detail that philosophers go into, and I felt that they often missed the forest for the trees. I just wanted to paint bold ideas. But I eventually came to see that this sort of rigorous detail is vital for getting clear on the issues. And I found that the foundational issues in metaphysics and the philosophy of language that I had initially neglected were crucial to understanding consciousness, and also beautiful in their own right. I hope my younger self doesn't see me as having been brainwashed!

Some part of me still self-identifies as a scientist. And while doing philosophical work, I've also wanted to engage the issues at a level that people outside philosophy can understand, and in particular in a way that resonates with people in science. So I've ended up pursuing these issues on two fronts. Some of the things I write are technical and aimed mostly at philosophers, and other things are much more general. It can be awkward when the two audiences mix. A philosopher might find the general work unsophisticated, and scientists are often bemused by esoteric talk of zombies, supervenience, and possible worlds. But I think working at both levels helps keep you honest. As Kant said, big ideas without details are empty; details without big ideas are blind.

Chrucky: You argue in your work that neuroscience will not be able to give a complete theory of consciousness. Do you think that current scientific work on consciousness is misguided?

Chalmers: Sometimes the sort of non-materialist view I put forward is seen as anti-scientific, but I don't see it that way at all. I argue that neuroscience alone isn't enough to explain consciousness, but I think it will be a major part of an eventual theory. We just need to add something else, some new fundamental principles, to bridge the gap between neuroscience and subjective experience. Actually, I think my view is compatible with much of the work going on now in neuroscience and psychology, where people are studying the relationship of consciousness to neural and cognitive processes without really trying to reduce it to those processes. We are just getting much more detailed knowledge of the associations and correlations between them. Things are still in early stages, but one can imagine that as we build up and systematize our theories of these associations, and try to boil them down to their core, the result might point us toward the sort of fundamental principles I advocate. Of course that's a long way off yet.

There are areas in which current scientific practice might be expanded, though. A science of consciousness, as I see it, needs to make a systematic study of the relationship of third-person data concerning brain processes and the like and first-person data concerning conscious experience. One can see some current work as doing something like this, but it's striking that the methods we have for characterizing the first-person data are very primitive compared to those involved on the third-person side. I think what's missing is a formalism for describing and taxonomizing the many aspects of conscious experience itself. There have been attempts in this direction over the years, in both philosophy and psychology, with mixed results, and it's not obvious that there will ever be a perfect formalism. But I think that we at least need to move in that direction in order to develop a sophisticated science that takes consciousness seriously.

Anyway, there is a lot of really interesting work going on in the neuroscience and psychology of consciousness, and I would love to see philosophers become more closely involved with this. For example, there are some fascinating experimental results with monkeys, aimed at isolating the "neural correlate" of consciousness. These raise deep methodological and conceptual

questions. How do you justify criteria for ascribing consciousness, given that we can't detect it directly? What does it mean, exactly, for a given system to be a "neural correlate of consciousness"? I've been working on this sort of question lately, trying to provide constructive analyses that engage the science in detail. In general I think that philosophers have a lot to contribute here.

Chrucky: Your name is being identified with the distinction between an easy and a hard problem of consciousness. Herbert Feigl made a related distinction in his long essay "[The Mental and the Physical](#)". In his "Postscript after Ten Years", he writes: "Some philosophers feel that the central issue of mind-body problems is that of intentionality (sapience); others see it in the problem of sentience; and still others in the puzzles of selfhood. Although I have focused my attention primarily on the sentience problem, I regard the others as equally important. But I must confess that, as before, the sapience and selfhood issues have vexed me less severely than those of sentience." He and his colleague, Wilfrid Sellars, couched it as the 'intentionality-body problem' (the easy problem) and the 'sensorium (or sentience)-body problem' (the hard problem). Did you formulate your distinction independently? Are you making the same sort of distinction in your problems?

Chalmers: There's certainly nothing original about the observation that conscious experience poses a hard problem. When I said this, I was stating the obvious. The context was a big interdisciplinary conference on consciousness, with people from all sorts of areas talking about the topics in all sorts of different ways. At this sort of event, one often finds people making big claims for a theory or model of "consciousness" that doesn't touch the deepest issues. The trouble is that the word "consciousness" is deeply ambiguous, so there's not just one problem of consciousness. So I found it useful, for purposes of clarity, to make the obvious distinction between the "easy" problems of consciousness, which involve such functional matters as discrimination, integration, and verbal report, and the "hard" problem of subjective experience. I never expected this to catch on in the way it did! Of course similar observations have been made by any number of people, and the distinction is obvious to anyone who thinks about the subject a little. All I introduced was a catchy name. I'm glad it has caught on, though, because it's made it much harder for people to just ignore the problem in the way they often did before.

There's no question that Feigl recognized the difficulty of the problem. His problem of "sentience" is of course just the problem I'm concerned with. "The Mental and the Physical" is a marvelous and intricate piece of work which more people ought to read. He grapples with the problem very seriously, and tries his hardest to reconcile the existence of consciousness with what we know of science. Feigl is often assimilated with the early identity theorists such as Place and Smart, but I think this is a mistake. His view is much more radical. Rather than "reducing" consciousness to what we know of the physical, Feigl wants to flesh out our view of the physical so that it can accommodate consciousness. It's not unlike a view put forward by Russell a few decades earlier, on which consciousness provides the intrinsic nature of certain physical states that science characterizes only from the outside. Leopold Stubenberg has a nice paper distinguishing the "Australian" and "Austrian" versions of the identity theory. Although I'm Australian, I find myself much more in sympathy with the Austrian version!

Sellars is also someone who took the problem of consciousness very seriously. His critique of the "myth of the given" is sometimes cited as a sort of demolition of the intuitions that give rise to the problem, but it's clear that Sellars himself didn't see it this way. In his Carus lectures, late in his career, he does some serious metaphysics to try to deal with the problem of the "sensorium". Like Feigl, he seems to want to expand our view of the physical, though in a slightly different way. I think he might be best seen as a sort of emergentist, who holds that new fundamental "physical" principles come into play in systems of a certain complexity, and these principles may account for consciousness.

On the various problems defined by Feigl and Sellars, I think their "sentience" and "sensorium" problems map nicely onto the "hard" problem. It's not quite so clear that the "intentionality", "sapience", and "selfhood" problems map onto the "easy" problems. I define the easy problems in terms of functioning. Some aspects of intentionality, sapience, and selfhood might be explainable in terms of functioning, but there may be other parts that go beyond this, such as the phenomenological components of intentionality and the self, which I think are absolutely central. So these problems may have a foot in both the "hard" and "easy" camps.

Chrucky: One thing that puzzles me, as a student of the history of philosophy, is the widespread controversy about the existence of qualia. This is surprising because almost all philosophers prior to roughly 1950 accepted the existence of qualia -- they referred to them, of course, by other names: sense impressions, phantasms, ideas, sense data, *sensa*, *sensibilia* -- though the term 'qualia' was also used -- by C. I. Lewis, I believe. Am I correct in believing this? And what happened to produce this widespread aversion to sense data or qualia?

Chalmers: Actually, I think most people accept the existence of qualia. Only a few deny them -- Dan Dennett, for example, and even there the denial isn't unequivocal. What's controversial about my own view is not so much that I defend the existence of qualia, but that I argue that they are nonphysical. Many contemporary philosophers would like to have their cake and eat it too, by accepting qualia and holding that they are physical. That would be nice if it worked, but at the end of the day I think it just doesn't work.

Sense data are much more controversial than qualia, because they are associated with a controversial theory of perception -- that one perceives the world by perceiving one's sense-data, or something like that. They're also associated with some controversial views in epistemology. Much of the reaction against sense-data was a reaction to this sort of theory, rather than a reaction to the notion of qualia per se. It may be that some people are dubious about qualia because they see all this baggage packed into the notion, or other baggage such as the view that qualia are known incorrigibly, or that they have a certain nature. But none of this is essential to the core notion of qualia, which is just that qualia are the properties that characterize what it is like to be in a given mental state. It's pretty hard to deny that there are such properties, although of course some philosophers deny it anyway.

For my part, I often prefer not to use the term "qualia", partly because of all this baggage, and partly because it suggests a sort of abstruse problem invented by philosophers. I tend to just use the everyday terms "consciousness" and "experience" (or "subjective experience") myself, as that's what we are really talking about.

Chrucky: Describing qualia as nonphysical leaves room for a kind of distinction Sellars and Meehl made in their paper, "[The Concept of Emergence](#)", in which they distinguish two senses of 'physical': a 'physical-2' which refers to properties of matter prior to the emergence of life (the properties studied by physics), and 'physical-1' which refers to the properties of living things containing sentience. As I understand Sellars, anything which is physical must be in space, in time, and have causal properties. When you say that qualia are nonphysical, you could mean that they are 'physical-1'. Qualia would not be physical-1 (nor physical-2) if they were epiphenomena (i.e. qualia lacking causal powers). What is your position relative to this distinction?

Chalmers: Well, I take materialism, or physicalism, to be the thesis that the only fundamental properties and laws in nature are those characterized by physics: space, time, mass, charge, and so on, and the various laws governing them. To speak metaphorically, we might say that after God had created all of physics and had set up the boundary conditions, everything else came along for free. My central claim is that this is false. One needs further fundamental properties to accommodate consciousness -- experiential properties, or proto-experiential properties. This is roughly to say that one needs new fundamental properties which are not physical-2 in Sellars' sense.

Sellars is suggesting, in effect, that they could be "physical" in a broader sense. I think that issue is largely terminological, but if people want to use the terms that way, they can. What's important is to get the shape of the view right, not what to call it. Personally, I think using "physical" in this weaker sense (Sellars' 'physical-1') empties "physicalism" of much of its content. For example, consider the sort of view in quantum mechanics on which consciousness alone is exempt from the Schrodinger rules of wave evolution, and on which it somehow acts to collapse the wave function. This view might count as "physicalist" in Sellars' weaker sense, but I think it is best seen as a sort of dualism -- at least a property dualism, with two deeply different classes of properties entering into different sorts of laws. But if someone wants to call this sort of view physicalist, they can. I've always wanted to be a physicalist!

There are lots of subtle gradations of views in the vicinity, of course. Some of them are closer to "physicalism" than others. For example, a property dualism is closer than a Cartesian dualism with two kinds of substance. It may also make a difference

whether the new properties play a causal role with respect to the original microphysical properties. And there is the very interesting view on which the novel properties somehow make up the "inside" of the entities characterized by physics. This last view may well deserve to be called materialist in some sense. But the real point is that on any of these views, one needs to introduce experiential or proto-experiential properties as fundamental.

I see the options as falling into three classes. There is epiphenomenalism, on which the new properties don't play a causal role, so I suppose they can't be physical-1 in Sellars' sense. There is interactionism, on which they do play a causal role, so they might or might not be physical-1, depending on whether they are located in spacetime. And there is "panprotopsychism", the last view on which the novel properties are somehow inside the microphysical network from the start. I am perhaps most sympathetic with the last view, which is beautiful and elegant if the details can be worked out. But I have days when each of them seems attractive. It all depends on how a detailed fundamental theory shapes up, sometime in the future.

Chrucky: It seems that it is possible to classify various philosophers by their stance towards qualia. Using C. D. Broad's terminology (in [Mind and Its Place in Nature](#)), those who view qualia as delusive are the eliminativists; those who view them as reducible or identical to physical properties of inanimate life are reductionists, and those who claim that qualia are neither eliminable nor reducible, like yourself, are non-reductive or emergent materialists. Although there don't seem to be too many non-reductive materialists around, I am wondering how your version compares with those of others.

Chalmers: Broad is another marvelous under-read metaphysician. I like his [taxonomy of the seventeen fundamental positions](#) on the mind-body problem, with various symmetrical kinds of materialism, mentalism, dualism, and neutralism. Some of his seventeen don't have many contemporary proponents, but some do. He dismisses out of hand the sort of materialism that holds that qualia are "delusive", but nowadays people like Dennett come close to this view. Then there are many reductive materialists, of course. There is also a popular view he didn't really consider, on which mental facts are not deducible from physical descriptions, but on which the mental is somehow physical in any case. These people often call themselves "nonreductive materialists". That view is perhaps halfway between Broad's "reductive" and "emergent" materialism. Then there are a good number of dualists, and a few mentalists and neutralists.

My own view might well be seen as a version of his own favorite view, "emergent materialism", which has new fundamental principles governing the emergence of consciousness from the physical. I would prefer to call it a "property dualism" rather than a version of materialism, but that's mostly a terminological issue. I also have a good deal of sympathy with certain forms of what he calls "neutralism", on which for example there might be a network of "proto-experiential" properties underlying both the physical and the mental. This sort of view can be motivated by the considerations of Russell and Feigl that we discussed earlier, and it also fits nicely with the idea of thinking about physics and consciousness in terms of information theory, for example, which is a speculative idea that I have explored a little.

Chrucky: Am I correct in saying that you hold that the phenomena of physics, chemistry, and biology can be logically deduced from the laws of physics, but that the phenomenon of consciousness cannot? The first thesis seems to be in contradiction with Broad's position. [According to Broad](#), the properties of chemical compounds cannot be deduced from the properties of their components in isolation, or from their properties in other compounds. For example, the properties of water cannot be deduced from the properties of hydrogen or oxygen in isolation, or from their properties in other chemical compounds. Could you explain your position vis-a-vis Broad's belief.

Chalmers: That's approximately right. I hold that most phenomena can be logically deduced from the facts of physics. But the facts here should be taken to include more than the laws. They include for example the distribution of every particle and field throughout space and time. These facts probably can't all be deduced from the laws of physics, if only because of the role of boundary conditions and indeterminism, so we need to include them directly. Once we are given this sort of complete microphysical information, I think that Laplace's demon, who models the whole universe at lightning speed, could infer the facts of chemistry, the facts of biology, and so on. But he couldn't infer the facts of consciousness.

The position with respect to Broad is tricky. Broad held, in effect, that there are "emergent laws" governing the behavior of matter in certain configurations, and that these laws are irreducible. So when particles assemble themselves into certain chemical configurations, for example, the laws governing their evolution can't be inferred from the standard laws that govern those particles in isolation. I think that Broad was probably wrong about that. What we know now about physics and chemistry suggests that the laws of chemistry are probably a consequence of the laws of physics, although we certainly don't know all the details yet. But even if Broad was right about this, it might not threaten my claim. If he is right, the facts of chemistry aren't deducible from microphysical laws, but they might still be deducible from the totality of microphysical facts. The microphysical facts would tell us just how the particles are behaving when they are in certain configurations, for example, and from there one could figure out the behavior of various chemicals. Something similar holds for biology, I think.

So even accepting Broad's view of biology and chemistry, the situation there will be different from the situation involving consciousness. Where consciousness is involved, one doesn't just have new patterns of evolution of existing qualities. One has wholly new qualities involved. These two different sorts of "emergence" weren't always distinguished by the British emergentists, but I think they need to be kept separate.

Chrucky: You hold that the presence of consciousness does not supervene on the components of a structure, but only on their functional organization. So if some system of things, for example, a computer or the people of China (Ned Block's example), exemplified an appropriate functional structure -- regardless of the components of that structure -- they would possess consciousness. In other words, the emergence of qualia depends only on an appropriate functional structure. Is this correct?

Chalmers: Right. I don't think that consciousness can be logically deduced from either structure or function, but it is still closely correlated with these things. Any two physically identical systems in the actual world will have the same state of consciousness, as a matter of natural law. The question then is: what sort of physical factors matter? Is carbon-based biology required for consciousness, or can anything with the right organization be conscious? I hold that what matters is the functional organization. If a silicon system was set up so that its components interacted just like my neurons, it would be conscious just like me.

There are various reasons to think this. First, I can't see that neurons have anything special going for them that silicon doesn't. Second, I'm struck by the fact that the most direct correlations between consciousness and the brain can be cast at the level of information processing -- when we find a certain abstract structure in physical color processing, we also find a similar structure in our conscious color space. So that suggests that abstract organization is important. And third, I think one can argue that if one's neurons were to be substituted one-by-one with silicon chips, then one's consciousness would stay the same. The alternatives are that it gradually fades across the spectrum, or that it disappears all at once, and I don't think either of those are very plausible. Of course it's an open question, but I think that in a century or two, once we start actually doing this sort of replacement, most people will accept this sort of view.

The same would apply, in principle, if there were a hundred billion tiny people in your head instead of neurons or silicon chips, as in Block's example. As long as they are interacting in the right way, I think this sort of system should be able to support conscious experiences just like mine. It's counterintuitive, but after all it's counterintuitive that a gray mass of a hundred billion neurons should support consciousness. And again, I don't see why replacing neurons with tiny people one by one should change things.

Chrucky: In reading your exchange with Searle, in the [New York Review of Books](#) [Exchange; reply 2], I was not sure how to take his charge that you are, on the final analysis, a panpsychist. If you are, you are in good company with someone like Alfred North Whitehead. Can you formulate your understanding of panpsychism, and in what way you are or are not a panpsychist.

Chalmers: In my book on consciousness I discuss panpsychism fairly sympathetically, but I'm not committed to the idea. What I am committed to is the idea that experience is somehow fundamental to nature, and that there are fundamental laws connecting experience with the physical processes we know and love. One natural way for this to go would be for some sort of

experience, or proto-experience, to be associated with every physical process. That would make a theory particularly elegant and simple, and it may also help integrate experience inside the causal order, rather than having it dangle outside as a sort of epiphenomenon.

The trouble, of course, is that people think the idea is crazy. In the book I argue that it's not as crazy as people think it is. There needn't be anything like "minds" as we usually think of them at the fundamental level, for example -- I'm not suggesting that electrons are having deep thoughts about the protons they're revolving around! It's just some sort of very simple, primitive analog of experience, going all the way down. I think one needs to take the idea seriously -- after all, it may be a requirement for a theory of consciousness that it contains at least one crazy idea. But after saying all this, I should say that panpsychism is just one way that a fundamental theory might go, and I think the true shape of such a theory is very much an open question.

Chrucky: With the rapid rise of the Internet, we have also witnessed a very rapid rise in the number of philosophers using it. In fact the amount of material is proliferating at such a great rate that it is becoming hard to keep up with it. I know that you have invested a great deal of your time in developing a very important web site. Can you tell me how you view the Internet, the use of the Internet for philosophy, and your own role in managing a web site.

Chalmers: I find the Internet useful for all sorts of philosophical purposes. I use e-mail a lot, and have had many productive philosophical discussions this way. And there is a lot of interesting material out there on the web. My own web site started as an experiment in my spare time, but it has mushroomed, and now it plays number of different roles for me. For a start, I have made a number of my papers available online, which makes them available much more widely than if they are buried in a philosophy journal somewhere. I am sure that a number of my papers have had many more readers through the electronic version than through the print version, especially among nonphilosophers. It's also useful for making unpublished material available, and for getting feedback. And I have recently started using my web site for posting brief responses to papers that discuss my work. Often one doesn't want to go do the length of publishing whole articles in response, but it's nice to be able to at least make some sort of response available to those who are interested.

I've also used my web site to make a number of resources available for the wider community. I started compiling an annotated bibliography of the philosophy of mind when I was a graduate student learning the ropes. People told me I should make it available to others, so I did, and now it has grown to more than 4000 entries. The web has made it much easier to use. I've also put together various pages with links to high-quality philosophy available elsewhere on the web. For example, it turns out that around 400 papers on consciousness are available online, many of them by the leading philosophers and scientists in the field. I put together a page categorizing these papers and linking to them. That page now gets a lot of use by people inside and outside the field, and it is growing all the time. On the lighter side, I have a page devoted to zombies, and one devoted to philosophical humor. Last time I checked, the humor page got the most traffic of any of my pages. I tend to work on my web page when I'm in a mood to avoid getting real writing done -- but at least it's a form of productive procrastination!

One effect of all this, I hope, is to make real philosophy much more accessible to people outside the field, and people outside academia. I have had a lot of nice feedback from people who appreciate this. Another effect is to help open lines of communication between the different fields involved in a broad area such as consciousness research. And a third effect is to speed up the process of philosophical dialectic. New work is often made available much more quickly these days. Where responses between philosophers might have previously taken years, at least at the public level, now they can be available in months. I suppose one negative effect could be potentially to encourage sloppiness and a lowering of standards, but I think that the existence of journals and such will keep standards high. Overall, I hope the effect of the Internet is to help make philosophy an even more interesting, active, and open field than it is already.