

Philosophy of Mind: An Overview

Laura Weed

In the twentieth century philosophy of mind became one of the central areas of philosophy in the English-speaking world, and so it remains. Questions such as the relationship between mind and brain, the nature of consciousness, and how we perceive the world, have come to be seen as crucial in understanding the world. These days, the predominant position in philosophy of mind aims at equating mental phenomena with operations of the brain, and explaining them all in scientific terms. Sometimes this project is called 'cognitive science', and it carries the implicit assumption that cognition occurs in computers as well as in human and animal brains, and can be studied equally well in each of these three forms.

Before the mid-twentieth century, for a long time the dominant philosophical view of the mind was that put forward by Ren é Descartes (1596-1650). According to Descartes, each of us consists of a material body subject to the normal laws of physics, and an immaterial mind, which is not. This dual nature gives Descartes' theory its name: Cartesian Dualism. Although immaterial, the mind causes actions of the body, through the brain, and perceptions are fed to the mind from the body. Descartes thought this interaction between mind and body takes place in the part of the brain we call the pineal gland. However, he didn't clarify *how* a completely non-physical mind could have a causal effect on the physical brain, or *vice versa*, and this was one of the problems that eventually led to dissatisfaction with his theory.

In the early twentieth century three strands of thought arose out of developments in psychology and philosophy which would come together to lead to Cartesian Dualism being challenged, then abandoned. These were Behaviorism, Scientific Reductionism and Vienna Circle Verificationism. I will begin with a very brief summary of each of those positions before I describe various contemporary views that have evolved from them:

Behaviorism: Behaviorists accept psychologist B.F. Skinner's claims that mental events can be reduced to stimulus-response pairs, and that descriptions of observable behavior are the only adequate, scientific way to describe mental behavior. So, for behaviorists, all talk about mental events – images, feelings, dreams, desires, and so on – is really either a reference to a behavioral disposition or it is meaningless. Behaviorists claim that only descriptions of objectively observable behavior can be scientific. Introspection is a meaningless process that cannot yield anything, much less a 'mind' as a product, and all human 'mental' life that is worth counting as real occurs as an objectively observable form of behavior. Head-scratching is objectively observable. Incestuous desire is not; nor is universal doubt, apprehension of infinity, or Cartesian introspection. Philosophers like Carl Hempel and Gilbert Ryle shared the view that all genuine problems are scientific problems.

Verificationism was a criterion of meaning for language formulated by the Logical Positivists of the Vienna Circle, who argued that any proposition that was not an a logical truth or which could not be tested was literally meaningless. For example, a mother's claim that the cat will bite Jimmy if he doesn't stop teasing her is testable, but a theologian's claim that the Infinite Absolute is invisibly bestowing grace in the world is not.

Scientific Reductionism is the claim that explanations in terms of ordinary language, or sciences such as psychology, physiology, biology, or chemistry, are reducible to explanations at a simpler level – ultimately to explanations at the level of physics. Some (but not all) mental terms can be 'operationalized', or reduced to testable and measurable descriptions. Only these ones will rate as real mental events to the scientific reductionist. There will be no Cartesian or Platonic 'mind' left over to be something different from a body.

Mental Events are Physical

The Oxford philosopher Gilbert Ryle (1900-1976) had another way to explain away the mind that Plato and Descartes believed exists independently of a body. Ryle characterized Cartesian

Dualism as a 'category mistake'. Category mistakes, as the name suggests, involve putting something into the wrong logical category. In Ryle's example, a visitor to Oxford wanders around the various colleges, libraries, laboratories and faculty offices, and then asks: "Can I see the University?" She has missed the fact that in seeing the buildings she was already seeing the university.

Ryle claimed that Descartes' 'ghost in the machine' – his immaterial mind in a material body – is a similar mistake. Descartes thinks he must have a 'soul' in his body that possesses his talents, memories and character. Ryle says that like the university, the mind is just the organization of Descartes' body's propensities. Bodies don't need a ghost to run them. According to Ryle, the properties of a person are better understood as adjectives modifying a body, than as a noun (an object) parallel to it. Intelligence, for example, is not a thing that exists apart from and parallel to a body, but rather is a collection of properties a body has. Intelligence includes properties such as social skill, quick wit, organizational ability, math ability, a sense of humor, musical talent, articulateness, critical thinking skill, and artistic sensitivity. Someone who never exhibited any of these skills or abilities would not be called intelligent; and anyone who is considered intelligent exhibits some of these talents.

Ludwig Wittgenstein (1889-1951) contributed an argument against private language. He claimed that for a symbol or a word to have a meaning there must be agreement among people about what the symbol is to mean. Plato's idea of what 'triangle' means was that an 'inner' mental image occurs 'in your mind'. By contrast, according to Wittgenstein, 'triangle' is a public word, used to communicate in a social group. Children learn its correct and incorrect applications by being corrected by elders in their use of the word. According to Wittgenstein, apart from the social use, 'triangle' has no meaning. Similarly on this account, 'mind' has no meaning apart from its effects.

J.J.C. Smart added materialism to scientific reductionism in this developing point of view by claiming that mental states could literally be particular states of the brain – so that for example

some C-fibres firing in one's brain would be identical with a specific feeling of pain. This became known as the Mind-Brain Identity Theory, and for a while it dominated philosophical discussions about mental events. Since then, however, Identity Theory discussions have been superseded by discussions driven by computer metaphors, such as Functionalism, Neurological Reductivist Materialism, Supervenience Theories, and Naturalistic Dualism. So let's look at those newer theories.

Functionalism

Functionalism is the theory that the important thing about mental states is not where they are located or what they are made of, but what function they perform.

Alan Turing is generally regarded as one of the fathers of computer science: among other achievements, he produced the first ever design for a stored-program computer. He also argued that artificial intelligence is intelligence in every sense of the word. In a 1950 paper he described a scenario which has since become known as the Turing Test. Suppose you are communicating with two people on the other side of a wall. You pass notes through a slot and figure out which of the people is responding to your notes. Now, suppose that one of the people is replaced by a computer, and you can't tell that this has happened. Do you have any reason to say that the person you were communicating with before is intelligent but the computer is not? Turing says, no, you don't. If intelligence consists of your ability to solve math problems, keep track of lots of information, organize data, recognize recurring patterns, and play chess, and the computer can do all of these things better and faster than you can, then you have no right to claim that you are intelligent and it is not. Now that Big Blue has beaten Kasparov at chess, and the best *Jeopardy* players have been beaten by IBM's Watson, Turing's claim seems even more convincing.

Turing is identifying mental properties with mental functions – not with observable behavior, as Ryle did; nor with brain states, as Smart did. Turing assumes mental functions can cause behavior and brain states, but not that they're identical with either behavior or brain states.

Hilary Putnam, writing in the 1970s, argued that a feeling of pain could be a function that is in principle realizable in a collection of silicon chips or some other physical apparatus as well as in a brain. Putnam called the idea that humans can think but computers can't, 'hydrocarbon chauvinism'. He further claimed that any organism can be described as a *probabilistic automaton* – i.e., as a system that undergoes transitions from initial states, through processing functions, to output states which can be predicted with varying degrees of accuracy. All organisms are systems that causally interact with the environment, have processing procedures, and output effects, claimed Putnam. (He has since changed his mind about functionalism and become a pragmatist.)

Jerry Fodor added to the functionalist program the proviso that any function capable of working as brain states do must be computational. Neurons, structures and patterns in the brain can be described in terms of mathematical models. Therefore if mental events are to be functionally connected to brains in a one-to-one correspondence, then they too must be realizable through a language of thought in a digitizable format.

Neurological Reductionism

Paul and Patricia Churchland espouse a position they call 'eliminative materialism', which argues that the project of neuroscience will actually prove to be even more radical than identity theorists like Smart realized. The Churchlands claim that talk of mental states will eventually be abandoned altogether, in favor of a radically different view of how the brain works not identified with brain states.

According to the Churchlands, folk psychology is the way most people think about how thinking works. So for example, most people now think that we have a stream of consciousness that contains images and conceptions of a wide variety of types about which we have beliefs and attitudes. Our beliefs and attitudes are colored by our feelings, which include mental states like joy, sorrow, resentment, anxiety and relief. We also think that the way we sense the world and ourselves is largely a direct representation of the way the world is; so the world contains

cold and hot, colored, shaped, hard and soft, threatening and soothing things, and our bodies sometimes are those ways as well. All of this is false, according to the Churchlands. It is not just a bit misleading, the way a fuzzy map might misrepresent some areas of terrain. It is downright false across the board, in the way that the notion that demonic possession explains mental illness is false.

Paul Churchland points out how radically scientific revolutions alter the way people think about things. When Aristotle's theories in physics were replaced by Newtonian physics, his ideas like 'natural motion is circular' just ceased to exist. Likewise, science now has no place for phlogiston, choleric personalities, and demonic possession. Churchland predicts that in the same way, at some point in the near future, people will no longer even try to introspect to see how they are doing. Just as a psychologist might now tell a depressed patient to stop worrying about why he is depressed and take some Prozac, so in the future, people might figure out how they are doing mentally by giving themselves a home fMRI or CAT scan and having their computer analyze the data. The resulting analysis will have nothing in common with "I'm sad because my cat died," or "I'm elated over the beautiful sunset."

Churchland has three arguments in favor of eliminative materialism. The first is that folk psychology fails to explain such common activities as sleep, learning, intelligence and mental illness. Since folk psychology has been around for thousands of years, it isn't lack of time to work out the details folk psychology suffers from, it is explanatory poverty. Secondly, the history of ideas supports elimination of old conceptual frameworks. Folk notions of motion were completely replaced by Newtonian physics, leaving not a trace. Folk ideas of cosmology, fire and life were equally cockeyed. The phenomena of conscious intelligence are more complex and harder to understand than any of the above, so there is little likelihood that our folk ideas about consciousness could be right. Thirdly, it is highly improbable that folk psychology will be reduced to neurobiology. Reductions require that the specific principles and types of things in one theory closely mirror those in the reduced theory. Neurobiology is highly unlikely to do this.

Daniel Dennett adds to the Churchlands' project a claim that interpreting a system as an intentional and rational system is simply a matter of taking a particular type of stance with respect to the system. To see Big Blue the chess computer as rational and interpret its movement as planning to attack a queen, is simply an admission that we don't know what design or physical features of Big Blue produced the behavior we observe, and so the behavior appears rational. Complex systems, says Dennett, appear intentional when viewed 'from the top down', and mechanical when viewed 'from the bottom up'. To Dennett, agents, intentionality, meanings in language, phenomenal qualities, intelligence in the abstract, and mental entities in general, can play no engineering role in explaining the workings of any system, human or otherwise. So, in all cases of apparent rationality, apparent agents can be decomposed into mechanical parts.

Supervenience

Donald Davidson and Jaegwon Kim agree with the reductionists that only physical and mechanical principles explain anything. But they insist that phenomenal experience, such as the experience of seeing a sunset, adds something to a human life that a computer might lack. Kim and Davidson both said that phenomenal qualities are *supervenient properties* of brains: properties arising simply because the physical processes in the brains were working. The supervenience of mental phenomena on brain activity like this is understood as paralleling the supervenience of smoke on fires: the smoke does not causally effect the fire, but will be there, as a by-product, whenever a fire is occurring. These philosophers thus avoided denying the reality of mental experience, but the supervenient phenomenal properties are here viewed as playing no causal role in thinking or action. This supervenient view, of mental phenomena being causally-ineffective emergent properties of the brain, is similar to the position in philosophy of mind called *epiphenomenalism*.

Naturalistic Dualism and the Hard Problem

David Chalmers, however, argues that materialist reductionism of the Churchlands' type throws out too much, and cannot deal

with the fact that humans enjoy sunsets. Chalmers agrees with Thomas Nagel that there is something that it feels like to be a bat, or a human, but there may be nothing that it is like to be a TV set. (Computers are left an open question).

Chalmers argues that functionalists and reductionists are only dealing with the 'easy problems' of consciousness. Problems such as how an organism learns, how the sense mechanisms work, how the brain processes sensory input and the like, are all mechanical questions about organic functions, so as one would expect, mechanical explanations are adequate to explain them. The hard problem, according to Chalmers, is why any of these events should be accompanied by phenomenal experience: what it's like to see red, for example. He argues that there are no physical facts about brains from which it follows that phenomenal experience should occur for those and only those physical events for which it does occur. In other words, there's nothing physically special about the brain which explains experiences. Further, rejecting behaviorism, Chalmers points out that a first-person perspective is required to even know that phenomenal properties accompany the physical events.

Chalmers argues for a form of dualism that he calls 'naturalistic dualism'. To explain consciousness in full, he argues, requires taking phenomenal experience seriously. But, unlike Plato and Descartes, Chalmers believes that the conscious phenomena are dependent on the existence of brain states. This implies that the relationship between the mental states and their biochemical base is scientifically discoverable. Also, the conscious states must mirror the functions performed by the biochemical states in some important ways. Chalmers also calls his position 'non-reductive functionalism'.

Objections to the Cognitive Science Program

While John Searle agrees with the materialist leanings of the cognitive scientists, he has been arguing that functionalists and eliminativists take the computer model too seriously, as actually descriptive of the functioning of a mind (Strong Artificial Intelligence) rather than as a helpful metaphor (Weak AI). Searle's two main objections to Strong AI concern the distinction

between syntax and semantics in language, and the distinction between causation and logical inference in reality.

The syntax of a sentence is the grammar or logical structure of the sentence. It can be captured through a formulation of this structure in symbolic logic. The semantics of a sentence is its meaning or reference. Searle says that philosophers like Turing, Fodor, the early Putnam and other advocates of Strong AI collapse semantics into syntax. There are some reasons for doing this. For instance, Turing could translate the code that the Germans were using in World War II using only his syntactic engine, without reference to the meaning of what he was translating. However Searle argues that in the way they operate, languages do not collapse semantics into syntax. He makes this point most clearly through his Chinese Room example. A person who speaks no Chinese, sitting in a room, has cards with Chinese characters on slipped under the door to him. He has a rule-book for processing these characters, and passes further character cards out of the room according to those rules. A person outside the room interprets the output as someone answering questions in Chinese. Searle says that the ability to string Chinese symbols together according to grammatical or logical rules does not however constitute speaking Chinese, because the person in the room does not understand the reference or meanings of the symbols that a speaker of Chinese would give them. To understand the meanings, one would have to understand not only what the cards refer to, but a lot about Chinese culture, nuances of tone and context, social structure, mannerisms, etc. None of this data is contained in or reducible to the syntactical rules of Chinese.

Searle's second point concerns the distinction between causation and logical inference. Since the AI revolution began in the late twentieth century, a good deal of philosophical effort has gone into trying to show that a specific logical formula 'p implies q' is equivalent to or somehow reducible to the scientific claim 'p causes q'. Searle says there are several serious problems with this project. The main one is that logical relations are time-insensitive, and, for the most part, symmetrical: since 'p implies q' is equivalent to 'not-q implies not-p', I can derive either from

the other in either order. Yet causation is neither time-insensitive nor symmetrical in this way.

To Searle, the reason computational logic patterns can't be causal explanations of mind/brain behavior is that they are simulations. He points out that simulating a hurricane on a computer may tell you some things about the hurricane, but it doesn't constitute causing a hurricane. And the simulation has no causal power to make the hurricane do anything, such as change course or grow less powerful. Likewise, simulated fires don't burn anything, and simulated car crashes don't bend any metal. Simulated logical patterns don't cause mental states or influence brain states. Searle accuses the Strong AI people of confusing their virtual reality with the real thing.

Further Objections to Reductionism

Emotions: Recent discoveries by Antonio Damasio and Jaak Panksepp about the role of emotions in decision-making and social reasoning have raised further doubts about the strongly cognitive model of mind inherited from Descartes and perpetuated by the Strong AI /Turing machine model. Far from being the distractions to mental operations that Plato and Descartes represented them as being, emotions have turned out to be essential elements in mental functioning. Patients with pre-frontal-cortex brain injuries, like the railroad worker Phineas Gage [see [here](#)], or other brain injuries that impair emotional functioning, become incapable of even simple planning. Without emotional drive, cognition appears to become dysfunctional, at least in humans.

The Extended Mind: Other critics of the reductionistic agenda in the philosophy of mind have pointed out that many aspects of our mental functioning are not brain-bound in the way identity theorists supposed. The psychologist J.J. Gibson articulated the idea of human thinking as ecologically embedded in a body and an environment. Following this, Andy Clark argues that one's body, ability to move, and system of environmental affordances, are as much a part of one's mental functioning as are brain functions. Clark shifts the philosophical emphasis from analysis of the brain to analysis of a human's kinesthetic interaction with

an ecological and social space. He points out that large-scale social projects, such as a building project or a disaster relief effort, occur across a considerably extended space and through the intersection of many people's minds, and are not limited to neuronal firings in any individual brain. Clark, in a joint paper with David Chalmers, discusses the fictional example of Otto, a man with memory problems who remembers the location of a library (and other useful pieces of information) by writing it down in a notebook. They argue that Otto's memory is literally in the notebook, not in his brain. Similarly, much of the memory of all of us arguably now resides in a variety of electronic devices.

Panpsychism: A more robust form of criticism of the reductionist program comes from a revival of panpsychism by philosophers such as Galen Strawson and Gregg Rosenberg, and physicists such as Henry Stapp. They concur with Alfred North Whitehead's view that for consciousness to be *anywhere* in nature it must be *everywhere* in nature, and with William James' view that our stream of consciousness is open to intrusions from an environmentally-pervasive conscious 'more'. In other words, everything has an element of consciousness. For most of the materialists, consciousness exists only as a rare occurrence in the brains of a single or a few species (if at all). The panpsychists charge that on this account, consciousness is a complete 'ontological dangler': a few anomalous islands of consciousness surface, for little apparent reason, in a vast sea of insentient and unconscious dead matter. Strawson, Stapp and Rosenberg object that the materialist picture arises from a Newtonian misunderstanding of matter. However, in quantum physics, matter may not be insentient, unconscious and dead, but have an element of consciousness too.

© Prof. Laura Weed 2011

Laura Weed is Professor of Philosophy at the College of Saint Rose in Albany, New York.

Here are some handy hints for further reading:

José Luis Bermudez, Anthony Marcel & Naomi Eilan, eds., *The Body and the Self*, 1998

David Chalmers, *The Conscious Mind*, 1996

Paul Churchland, *Neurophilosophy*, 1986

Andy Clark, *Supersizing the Mind*, 2008
 Antonio Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*, 2003
 Donald Davidson, *Essays on Actions & Events*, 1980
 Daniel C. Dennett, *Consciousness Explained*, 1991
 Jerry Fodor, *The Language of Thought*, 1975
 Jay Garfield, *Foundations of Cognitive Science: The Essential Readings*, 1990
 Thomas Nagel, *The View From Nowhere*, 1986
 Jaak Panksepp, *The Archaeology of Mind: Neural Origins of Human Emotion*, 2010
 Hilary Putnam, *Mind, Language & Reality*, 1975
 Gregg Rosenberg, *A Place for Consciousness*, 2004
 Gilbert Ryle, *The Concept of Mind*, 1949
 John Searle, *Intentionality*, 1983
 J.J.C. Smart, 'Sensations and Brain Processes', *Philosophical Review*, vol. LXVII, 1959
 Henry P. Stapp, *Mind, Matter and Quantum Mechanics*, 2nd edition, 2004
 Galen Strawson, 'Realistic Monism: Why physicalism entails panpsychism', *Journal of Consciousness Studies*, 13, 2006