In Defense of Teleological Behaviorism

Howard Rachlin
Stony Brook University

Teleological behaviorism (TEB) is a behavioral identity theory that denies what seems to be the intuitively obvious notion that our minds are contained in our bodies—more specifically, in our brains. TEB holds, instead, that our minds are identical with abstract, temporally, and socially extended patterns in our overt behavior. TEB depends not on efficient causes but on a wide (or inclusive) version of final causes. This article illustrates how TEB accounts for mental acts and defends it against other theories of mind such as neural identity theory and double-aspect theories.

Public Significance
This article presents and defends a way of understanding the human mind, different from both common sense and modern psychology and philosophy. This view may lead to advances in the understanding of behavior (such as self-control) normally attributed to internal causes.

Keywords: behavioral identity theory, consciousness, dualism, efficient causes, final causes, mentalism, neural identity theory, teleological behaviorism

Teleological behaviorism (TEB) is a behavioral identity theory of mind. It identifies the mind not with internal events but with overt patterns of behavior observable in principle by other people. Overt behavioral patterns are said by TEB to cause the particular acts that comprise them. Thus, according to TEB, the mind may cause overt behavior. However, TEB’s notion of cause differs from our usual notion; for TEB, following Aristotle (see Rachlin, 1992, 1994, 2000), the mind may be a final cause of behavior (an answer to the question, WHY?) but not an efficient cause of behavior (an answer to the question, HOW?). Of course, to explain behavior fully, one must have answers to both questions. But mental causes, the domain of psychology, are, for TEB, final causes. Thus, the notion of final causation is crucial to TEB.

The history of science is full of prima facie facts proven to be less useful than their contraries. Especially suspicious are those facts that put us at the center of the universe (physical or spiritual). From a pragmatic viewpoint (that of TEB), something is true because it is useful in the long run to behave as if it were true. The burden is on TEB to show that this account is more useful than others. Once that happens, what seems obvious will change accordingly. It is perhaps fair to say that there has not yet been enough understanding of TEB to give it a fair test. The purpose of this article is to correct this lack.

Final Causes
It is possible to distinguish between two kinds of final causes: narrow final causes and wide final causes (the philosopher, Ackrill, 1980, calls them “dominant” and “inclusive” causes). Let us consider the narrow (or dominant) kind first.

Narrow Final Causes
Narrow final causes are congruent with the behavioral concept, “reinforcement”; when a
high-valued act (such as a rat eating a food pellet) is made contingent on a lower-valued act (such as pressing a lever), eating the food pellet is said to reinforce the lever press (Premack, 1965). For teleological behaviorism, eating the pellet is the narrow final cause of lever pressing. Q. Why did the rat press the lever? A. To eat the food pellet. Suppose that you put a dollar into a candy machine, press a button, and get a candy bar. The candy bar is the narrow final cause of the sequence: put dollar in slot—press button. Putting the dollar in the slot, pressing the button, and eating the candy are separate acts. Narrow final causes act in the opposite direction to efficient causes. Inserting the dollar and pressing the button cause the candy to appear in the tray (efficient cause). And, the future appearance of the candy causes you to insert the dollar (final cause). Q. Why did you put the dollar in the slot? A. To get the candy.

Wide Final Causes

Wide (or inclusive) final causes describe the relation of abstract, temporally extended patterns of activity to the particular acts comprising those patterns; the pattern is said to be the wide final cause of the act. For example, performing a dance is a wide final cause of doing the steps of the dance. Q. Why are you doing those steps? A. Because I am doing that dance. Playing a sonata is a wide final cause of playing the notes of the sonata. Why are you playing those notes?—because I am playing that sonata. Why are you pitching, or catching, or running the bases?—because I am playing baseball. The pattern is the cause of the particular act, and the act is the effect of the pattern. All the notes of a symphony (particular acts) must be played before the symphony (their pattern) can be said to be played, and in that sense particular acts precede their wide final causes, but it would be more accurate to think of wide final causes as embracing their effects than as following their effects (as narrow final causes do). Just as efficient causes may form a chain, with the effect of one act causing another act, so wide final causes form nested sequences; relatively particular acts are nested within relatively abstract acts like a set of Russian dolls.

Imagine that you saw a snippet of film of a man swinging a hammer. You do not know why he is swinging it. He could be aiming at some-one’s head. But then you see more of the film and he’s hammering a nail. Why is he swinging the hammer?—to hammer a nail. Then you see still more of the film, and you see he’s joining one piece of wood to another. Why is he hammering the nail?—to join one piece of wood to another. You see more and more of the film, and you see he’s building a floor, building a house, providing shelter for his family, supporting his family, being a good husband and father, until finally (you would have now have to see a film of virtually his whole life) being a good person. Each more abstract pattern is a wide final cause of all of the more particular acts. Each more particular act is an effect of all of the more abstract ones. Playing Beethoven’s Fifth Symphony is the final cause of playing each movement. Being a good person is a final cause of being a good husband, a good father, a good friend, and so forth. Just as particular acts may have several efficient causes, so particular acts may be part of more than one pattern. For example, a single note in a Bach fugue may be part of two or three or more overlapping themes, and a single act, such as a person crossing the street, may be, at the very same time, part of the patterns of shopping, exercising, visiting a friend, and so forth.1

From the wide view, a dance is a final cause (but not the only one) of a step; a program of dances is a final cause of a specific dance; a dance career is a final cause of a dance program; a happy life is a final cause of a dance career. Each more abstract and temporally extended pattern is a final cause of its narrower component. A final cause, such as dancing a dance or playing a game of baseball or playing a sonata on the piano, may be said to embrace its particular effects (steps, shots, themes). From the wide view, the relationship between a cause and its effect need not be 1:1. For example, living a healthy life may be a final cause of regular exercise, but no particular positive consequence may follow from each push-up. I believe that the wide view is the more useful one (Rachlin, 1994, 2000). From now

---

1 Similarly, for TEB, my trip to the cash machine may be part of two patterns: my desire to get $200 and my knowledge that I have that much in the bank. Aristotle (De Anima, book III, chap. 3, 429a, 1) compares such overlapping of patterns to points of tangency between two circles. That is, a single point is part of both circles as a single act is part of both patterns.
on, when I speak of final causes, I will be taking the wide view.

Efficient causes are followed by their effects; wide final causes are comprised of their effects as dances are comprised of steps or melodies comprised of notes. According to TEB, therefore, the mind (in the form of an abstract pattern of overt acts) may be the cause of a particular overt act. Patterns of overt acts such as wishes, hopes, intentions, perceptions, imaginations, and so forth can thus cause particular acts. Of course, abstract patterns would not exist without their particulars; a dance, to exist, requires steps. Nevertheless, abstract behavioral patterns, such as dances, do not exist in another world or only in our heads. They exist in our world and are no less real than the particular acts of which they are composed. Even a step, although particular relative to a dance, is an abstract pattern relative to the individual muscle movements that make it up. That does not mean that a step is any less real than a muscle movement.

Final causes are not necessarily any less scientific, less accurate, or even less precise than efficient causes. In physics, Maxwell’s equations (and all of field theory) work with final causes. According to Max Planck, a founder of quantum theory,

The \textit{cause efficiens}, which operates from the present into the future and makes future situations appear as determined by earlier ones, is joined by the \textit{cause finalis} for which, inversely, the future—namely a definite goal—serves as the premise from which there can be deduced the development of the processes which lead to this goal. (Yourgrau & Mandelstam, 1968, p. 165)

Final causes are also fundamental in thermodynamics and in economics. The discount functions (delay, probability, and social), prevalent in modern behavioral psychology and behavioral economics, are also final causes—unless one (erroneously) refines them as neural states inside the head. Final causes are particularly suited to psychology. Self-control and social cooperation may better be seen as conflicts between abstract behavioral patterns (healthy or moral patterns) and particular acts (impulsive or selfish acts) rather than as conflicts between inner forces such as “willpower” and outer temptations. I am not arguing that just because economics or modern physics makes use of final causes, psychology should do so too. I am saying that, because wide final causes are relatively abstract, a science using wide final causes will not necessarily be less accurate or less precise than one based on relatively particular efficient causes.

**Mental Terms**

It is beyond the scope of this article to provide a dictionary of mental terms from the viewpoint of TEB. As examples, the following two sections illustrate TEB’s approach to the terms \textit{perception} and \textit{imagination}.

**Perception**

For TEB, perception is identical to a correlation over time between a person’s overt behavior and an identifiable pattern of events in the environment. Consider the following question: What is the difference between two people (say John and Marcia), one of them (Marcia) stone deaf, both sitting stock still while a Mozart quartet is playing? A. John is hearing (i.e., perceiving) the music, whereas Marcia is not hearing it. Q. What does it mean to hear? A. To discriminate by overt acts, over a period of time, between sounds and silence. That is, a nonzero correlation exists between John’s behavior and sounds (unsignaled through other senses), whereas there is no correlation (a zero correlation) between Marcia’s behavior and such sounds. During the past, in the presence of sound signals, their behavior (perhaps including taking audiometric tests) differed and will differ in the future. (Consider their differing reactions to someone rushing into the room behind them yelling, “Fire!”) Their identical behavior during the Mozart quartet is merely one congruent point in two drastically different correlations between behavior and sound.

It could be that Marcia’s hearing mechanism is entirely normal but she is nevertheless unresponsive to sounds. In that case, we would say she was “psychologically deaf.” Is psychological deafness real deafness? From the viewpoint of TEB, the answer must be yes. What counts for deafness as for all psychological (or mental) states, for TEB, is Marcia’s behavior in the long run. If she was faking deafness, then her subsequent behavior would reveal what her state really was. If, despite her normal hearing mechanism, she continued to behave all her life as a
deaf person behaves, the question “Was she faking deafness or psychologically deaf?” would be entirely nonpragmatic—like whether a man and a squirrel chasing each other around a tree are or are not going around each other—and thus meaningless (James, 1907/1955, pp. 41–42).

**Imagination**

The teleological conception of imagination follows from that of perception. Aristotle says, “Imagination must be a movement resulting from an actual exercise of a power of sense” (*De Anima*, book III, chap. 3, 429a). As far as the overt speech and actions of a person are concerned, imagination is the same as perception. If I am doing a good job of imagining that I smell a rose, I will behave, for a moment, just as I would behave if I actually smelled a rose. The difference between perception and imagination is that the object is present in the world during perception (the rose is there when you are sensing it), whereas during imagination, the object is not present in the world (the rose is not there when you are imagining it). It is not necessary to infer that the rose I am imagining (that would be present in the world if I were perceiving it) is present inside me (as a representation, an internal image, a neural discharge, or anything else) when I am imagining a rose. When I imagine a rose, my overt movements with the rose absent are the same as those I would make if a real rose was present. In other words, all is the same in perception (or sensation) and imagination except that when I imagine the rose, it is not present.

If you generally behave one way in the presence of, and another way in the absence of, red lights, you are perceiving red lights. However, if, on occasion, you behave in the absence of a red light as you normally do in its presence, you are on that occasion imagining a red light. Imagining is acting and not dreaming: Vividness of imagination is not vividness of interior image but of overt behavior. Suppose two people are asked to imagine a lion present in the room. One closes her eyes and says, “Yes, I see it, a mane and a tail, that’s right, it’s walking around,” and so on. The other runs screaming for the door. You could say that they are both imagining a lion but in different ways. TEB puts it differently: Although they are both imagining, for TEB they are imagining different things. The first person is not imagining a lion but a picture of a lion. The second person is imagining the lion. The location, intensity, orientation, or even the existence of an image in the head of either of them would be entirely irrelevant to the imagination of either. And neuroscience bears this out. There are no pictures in the head for either of the two imaginers to look at and, even if there were, as Aristotle pointed out, there are no sensory receptors in the head to see them with. According to TEB, a good imagination is not just an aid or a tool in good acting. Rather, good acting is good imagining.

**TEB and Psychological Investigation**

Miller (1962), one of the founders of modern cognitive psychology, entitled his introductory book *Psychology: The Science of Mental Life*. TEB accepts this implied definition but proposes a particular way of studying mental life—not in terms of events within the brain but in terms of patterns of overt behavior. There are (at least) two ways by which mental events, such as a person’s intentions, may be understood. One way, that of modern cognitive psychology, is to observe the person’s behavior and infer from those observations what the inner mechanism must be to have given rise to that behavior. (This method is much like trying to infer the program of a computer by typing its keys and observing what appears on the screen.) Such an endeavor may be helped by using MRI machines to observe events actually going on inside the nervous system or by drawing an analogy to events measured directly in the brains of other species. Such cognitive experiments (e.g., Kahneman, Slovic, & Tversky, 1982) have improved our understanding of human behavior. But their ultimate purpose is to determine the actual mechanism by which the observations are generated. And that mechanism can best be described (they believe) by our mental vocabulary—hopes, fears, thoughts, and so forth acting as inner efficient causes.

Another way to study mental events, such as intentions, is by teleological analysis (Rachlin, 1992, 1994, 2014). This method is observation

---

2 All references to Aristotle are from McKeon (1941).

3 See Rachlin (1989) for a behavioral interpretation of these important experiments.
and analysis of patterns of behavior (including verbal behavior) over time. Mental terms, claims the teleological behaviorist, refer to these observable patterns; they exist on what the philosopher, Dennett (1978), calls the personal level. Consider how TEB approaches a typical mental term such as a person’s intention.

Suppose John has asked Mary out for a number of dates. (On one occasion, he tried to kiss her but she rebuffed him, hurting John’s feelings.) Mary wonders if John is now serious about the courtship. What is his intention? Is it just casual flirtation, or something long-term, or possibly marriage? For a teleological behaviorist, John’s past actions (including what he says to Mary and to other people) are the only relevant data on this question because it is in those actions where his intentions actually reside. Good data in this regard may be difficult to obtain, and John’s intentions may be obscure. But they are obscure, says the teleological behaviorist, because there is not enough currently available behavior to analyze, not because John’s intentions are hidden inside his head. The teleological behaviorist would never suppose that John’s true intentions could be revealed by a lie detector test that measures his brainwaves or his physiological responses to a series of questions, no matter how probing. However, John knows that he is trying to seduce Mary (let us assume) and the moment after he succeeds, he will up and go. Is this intention (or for that matter, his knowledge of his intention) in his head? No it is not, says TEB. His intention is in his past (and to some extent his future) behavior. The reason why John knows his own intention is not that he has access to something in his own head and Mary does not, but that John has access to his own past behavior and Mary does not. In principle, John’s twin sister, Jane, may know better than John does what his intentions truly are. When John tells Jane about his intentions, she may reply, “No you’re wrong. She’s got you in her grip with all that coyness, that playing hard to get. You’ll never leave her!” And Jane may be right. John’s belief, Mary’s belief, and John’s sister’s belief regarding John’s intentions are discriminations among complex patterns in John’s behavior. The expression of the belief at the present moment may be a simple assertion, but the discriminative stimuli on which the assertion is based (John’s behavioral patterns) are highly complex.

For TEB, mental terms are not just a loose way of talking about behavior, nor do they refer to covert events. For TEB, there are no inner psychological causes at all. It is not that the organism is in any way empty but that the molecular substrate of behavior (its set of inner efficient causes) is held by TEB to be the domain of neuroscience and not of psychology. Psychology should not rely on neuroscience but should stand on its own as a separate discipline. When some aspect of voluntary behavior is unexplained by current observation, a psychologist should look for its explanation, not more deeply in the nervous system but more broadly in time—in the behavioral history of the individual or, if still not found, in the evolutionary history of the species.

For Aristotle, understanding any particular act of an organism requires study of both its inner mechanisms (efficient causes) and the outer patterns into which it fits (final causes). TEB does not deny the importance of inner mechanisms—quite the contrary. But TEB views the study of inner mechanisms as the domain of neuroscience and it views the study of overt behavioral patterns as the domain of psychology. TEB could not and does not object to neuroscience; recent advances in this field have revolutionized our understanding of how the brain works. However, TEB does object to applying the vocabulary of psychology (i.e., mental terms) to inner events.

Nor could TEB object to the everyday-life use of mental vocabulary to refer to inner states. (In the great majority of everyday-life references to the relative motion of the Earth and the

---

4 According to Aristotle (Physics, book II, chap. 7, 198a), it is the business of physicists (including those whom we call psychologists) to know all the causes of their object of study. Aristotle’s concept of cause was much wider than ours. A cause, for him, was whatever could follow the word “because” in a sentence (Hocutt, 1974; Randall, 1960, p. 124). Aristotle discusses four types of causes: material, formal, efficient, and final. The former two refer to all existing substances. For instance, an object, a piece of cardboard (material cause), may have the form of a circle (formal cause). The efficient and final causes discussed in this article apply to movements of objects (i.e., behavior). For Aristotle, formal and final causes were more scientific, because more abstract and long-lasting, than material and efficient causes.
Sun, the Sun is said to be moving. Yet we all believe that in reality, the Earth is revolving.) One morning, we observe a piece of behavior—a person smiles. We say that he is happy, but we do not observe his happiness. We observe only the smile and harmlessly ascribe its cause to a state of happiness within his body. For most everyday-life purposes, this is sufficient. But, in reality, that smile is only one part of a pattern in his verbal and nonverbal behavior all morning. If that pattern is not there, he was not really happy (says TEB, and perhaps his wife and children) regardless of his inner neural state or his own introspective report.

**TEB and Philosophy**

The notion that our minds must be inside of us (let us call this notion “internalism”) is so firmly rooted in common sense that it may seem as if any other use of mental terms cannot possibly be called mentalistic. But, for many philosophers, mentalism is allied just as strongly with dualism as it is with internalism. These dualists (e.g., Block, 1981) cannot conceive how consciousness, for instance, can be purely physical. They attribute to monists, who say that the mind is part of the physical world, the implicit construction of homunculi as internal observers. Yet, despite common sense and common usage, and accusations of implicit internal homunculi, many if not most psychologists are monists and do believe that consciousness, and all of mental life, is physical. TEB is just one of several monistic, physical conceptions of the mind, which differ among themselves. I therefore ask the reader to suspend judgment regarding which if any of these theories are properly called mentalistic. Let us now turn to discussion of other theories of mind with regard to their approaches to the meaning of mental terms.

**Dualism**

Burgos (2015) presents an argument put forth by many modern philosophers of mind for dualism:

Zombies . . . are creatures physically exactly like us, down to the last particle. They thus have bodies and brains exactly like ours, anatomically and physiologically. They also behave exactly like we do, exhibiting pain and pleasure behavior, and reporting experiences of colors, shapes, smells, tastes, and everything we nonzombies report to experience. In short, zombies are physical twins of us nonzombies. Zombies differ only in lacking [nonphysical] conscious experiences: They feel no pain, no pleasure, have no sensory or perceptual experiences, even if their brains and behaviors exhibit everything we nonzombies do when we experience all of this. If zombies are logically possible, the argument goes, conscious experience cannot be physical. (p. 9)

To a teleological behaviorist, zombies are not logically possible. To say that zombies behave in every way (from birth to death) like a normal person with a mind, including crying in agony when tortured, and yet that they have no more conscious experience than a toaster is, for TEB, like saying A and not-A. For TEB, if an organism behaves consistently like a person who can feel pain, she can truly feel pain and is in pain when she so behaves. How do I know that at this very moment I am not the only nonzombie in the universe? How do I know that my wife is not a zombie? By the above conception of zombies, I do not. Any person I meet or any group of people may just as well be zombies.

Dualism of the kind where consciousness is said to be a nonphysical presence in the brain makes no sense because it provides no mechanism by which a nonphysical interior consciousness can learn about the outside world. Descartes himself realized this. The mind, in Descartes’s view, communicates with the world only through the nerves, which he conceived as the pulling of strings, the opening of valves, and the motion of animal spirits. But we do not see string-pullings, valve-openings, or the motions of animal spirits. We see chairs, tables, roses, lions, people. How do ideas of these things get into the mind? The answer, according to Descartes, is that God puts them there. We are born with all of the ideas that we will ever have. Stimulation of our sense organs (the pulling of strings) does not give us ideas, according to Descartes; it merely wakes up ideas that we already have:

Nothing reaches our mind from external objects through the sense organs except certain corporeal motions [the movements of animal spirits]. . . . But neither the motions themselves nor the figures arising from them are conceived by us exactly as they occur in the sense organs. . . . Hence it follows that the very ideas of the motions themselves and of the figures are innate in us. The ideas of pain, colors, sounds and the like must be all the more innate if, on the occasion of certain corporeal motions, our mind is to be capable of representing them to itself, for
For Descartes, all of our ideas, everything we know, is put into our nonphysical consciousness by God. This problem, of the transition from the physical world to the internal nonphysical mind and back again, is not avoided by modern dualists. According to them, our entire nervous system intervenes between the world and consciousness. Only in TEB, where the mind and the world are in direct contact, is there no such intervention.

Neural Identity Theory

According to the *Stanford Encyclopedia of Philosophy* (Smart, 2011),

The [neural] identity theory of mind holds that states and processes of the mind are identical to states and processes of the brain. . . . Consider an experience of pain, or of seeing something, or of having a mental image. The identity theory of mind is to the effect that these experiences just are brain processes, not merely correlated with brain processes.

Consciousness, according to this theory, would be reducible to neuroscience. A difference in the auditory cortex or still more centrally within the brain would be, according to neural identity theory, all there is to the difference between hearing and not hearing. Most neural identity theorists do not believe that hearing itself is just activity in a single nerve; for them, hearing is rather a pattern of activity over a complex of nerves extending over a wide area of the brain. Some neural identity theories would go beyond the brain to incorporate the peripheral nervous system. But, according to these theories, if we knew enough about the neural connections involved in audition, and we were able to build a brain that incorporated those connections, and if we were to feed into the artificial brain the strains of a Mozart quartet, that brain (which could not otherwise move, could not interact with the world, could not report its feelings, had no functions of use to anyone) would hear the Mozart quartet just as you or I (assuming you are not deaf) can hear it.

Certainly some particular pattern of stimulation in our brains is necessary for us to hear a Mozart quartet, and some movements of molecules in the air near our ears, and a particular pattern of movement of our basilar membranes, are necessary for us to hear the quartet. These are parts of the mechanism underlying hearing. But they are not the same thing as actually hearing any more than the activity of the engine of a car is the same thing as the acceleration of the car. Let me try to clarify exactly what the problem is here. You can say that a Mozart quartet is nothing but the playing of certain notes by four instruments. The music would be no less real or scientifically accessible for that. The quartet is abstract relative to its particular notes, but abstract entities are as real as particular ones. In certain senses of real, a chair is as real as (or more real than) the molecules that make it up. Neural identity theory is not problematic because it analyzes hearing and other mental events into neural events. If you produce the notes in their proper order and nuance, you will have all there is to the Mozart quartet. But, if you could produce the neural events said to be identical with hearing the quartet in a disembodied, artificially constructed brain, would that brain actually hear the quartet? Those of us who object to neural identity theory would answer this question in the negative. It is not conceivable to us that a disembodied brain could actually hear anything.

According to Rachlin (2005),

The problem with this neurocognitive identity theory . . . in the modern version, is that, despite advances in brain imaging techniques, there seems to be no set of physiological structures (molecular or molar) corresponding to each identifiable mental property. Modern sensory neurophysiology would not claim that a stimulus could be carried by a single chain of afferent nerves ending at a particular point in the brain. There are many stages on many brain levels as a stimulus affects the brain. Even a simple color is a diffuse pattern of nervous stimulation heavily influenced by contextual factors. Which [pattern] is identical with the sensation? If these factors came into focus at a “final common path” in some particular brain location you could call the focal point the sensation. But no single neural path for each sensation has been found within the nervous system. There is no place in the nervous system where the incoming stimulus stops. Sensory stimulation in its normal form runs right through the brain and out the other side so to speak without encountering anything identifiable as a sensation. (p. 46)

This is not to say that you could not build a robot with a human mind. However, such a robot would not need any particular human, internal structure. Rather, it would need to have a particularly human relation with human beings—to help people and be helped by them,
that is, to engage in economic interchange with them, to predict its own future behavior and theirs, to guide others, and also to mislead them (Rachlin, 2014).

**Double Aspect Theory**

Nagel (1998) and McGinn (2002) have presented double-aspect theories of mind: Body and mind are two aspects of the same thing. The question to ask two-aspect theorists is: Two aspects of what? Both Nagel and McGinn posit a third entity that the two aspects are aspects of. For Nagel, it is Factor X, and for McGinn, it is “unknowable.” Are these answers to the traditional question any more enlightening than Spinoza’s answer to that question—two aspects of Nature or God? I do not believe so.

To the question, “Two aspects of what?” Searle (2004) gives the (almost) correct answer. He says,

> Neural activity and conscious experience are different aspects, or levels of description, of the same thing, in the same way that, say, the molecular structure of a piston and the solidity of the piston are different aspects, or levels of description, of a piston. (p. 208)

Substitute behavioral activity (overt) for neural activity (covert) and this is TEB. But, if mental (or conscious) experience is an abstract version of physical activity, what is that physical activity? Why is it any more plausible for Searle, and the many philosophers who have considered this question, that conscious physical activity has to occur inside the head than that it occur in overt behavior? The word aspect implies a way of viewing (or “describing” something). Who views or describes the aspects of one’s neural activity? There is nothing in the brain capable of taking one or another view of neural behavior. On the other hand, different aspects of a person’s overt activity may be viewed or described by that person or by other people. Whole organisms, not their brains, have the sense organs needed to create that kind of description. Thus, according to TEB, individual consciousness and all mental activities are social constructions (Gergen, 1985).

Because Descartes believed that the soul interacted with the body deep in the brain and the physical motions had to directly influence the soul, and vice versa, the physical motions also had to be in the brain. But Searle presumably does not believe that there is a nonphysical soul present deep within the brain interacting with our nerves. Nor is this inherently obvious. Some societies and some ancient philosophers believed that our minds as well as our souls were in our hearts. Epicurus believed that all physical and mental actions are identical to the movement of atoms. He “maintains that soul atoms are particularly fine and are distributed throughout the body. . . . The human mind is . . . that part of the soul that is located in our chest” (Konstan, 2014). That is, Epicurus (as well as Aristotle and other Greek philosophers) located the mind elsewhere than in the brain. So it is not at all self-evident that the mind is in the brain or is the brain itself. This is an idea that we have been taught. It is possible to unlearn it. I would guess that if you name a vital organ, there will be or have been some society that believed it to be the location of the mind; there may even have been some societies that identified the mind with the whole organism. Descartes himself believed that the soul (even though it controlled the body through the brain) was coextensive with the whole body; thus, actions of the whole body were, for him, actions of the soul (which he did not distinguish from the mind). If the mind is a molar or abstract conception of some physical activity (as Searle seems to agree), and there is no a priori reason (such as connectivity with an internal, nonphysical soul) to assume that the physical activity occurs within the brain, then where does it occur?

In answering this question, usefulness is paramount, especially as consciousness, and talk of consciousness, must have evolved along with the rest of our human qualities. If our environment is seen as the source of our consciousness (as it would have to be if consciousness were a product of biological evolution), then it would be our overt behavior, not (as Descartes thought) neural behavior, which is in direct contact with the source.

There are strains of behaviorism in modern philosophy—for instance, the later Wittgenstein (1958) who said, “If one sees the behavior of a living thing, one sees its soul” (p. 357). Twenty-first century philosophy is not completely negative about behavioristic thought. Rowland Stout (2006) develops Ryle’s (1949) dispositional view of the mind. A disposition for Stout is a behavioral state of the organism. TEB’s view of the mind differs from Stout’s view in that for TEB mental terms refer not to immedi-
ate states but to temporally extended behavioral patterns.

According to Noë (2009),

Consciousness is not something that happens inside us. It is something we do or make. Better: it is something we achieve. Consciousness is more like dancing [overt behavior] than it is like digestion [covert behavior]. . . . The idea that the only genuinely scientific study of consciousness would be one that identifies consciousness with events in the nervous system is a bit of outdated reductionism. (p. xii)

After decades of concerted effort on the part of neuroscientists, psychologists, and philosophers, only one proposition about how the brain makes us conscious—how it gives rise to sensation, feeling, subjectivity—has emerged unchallenged: we do not have a clue. (p. xi)

The reason we do not have a clue is not because brain research is not important, or not interesting, or not valuable; brain research is highly important, highly interesting, and highly valuable. The reason we do not have a clue about how the brain makes us conscious is that looking for consciousness in the brain is looking for it in the wrong place.

Suppose that, at some time in the future, technology will have reached the stage where brains can be removed from human bodies and yet communicate wirelessly back and forth with bodies; the connection between the brain and the rest of the nervous system will be exactly the same as it is now except the connection will be “wireless” (neuronless) and will be applied in football to prevent concussions. The brains of both teams are stored safely in their locker rooms while the game goes on. Although there are no concussions, there are broken bones and lacerations, and crying out in pain, just as there are today. Now consider: Who or what is in pain? Clearly the players, not their brains, are in pain. Their brains bubbling away in the locker rooms are essential parts of the mechanism underling their pain just as the engine is an essential part of the mechanism underlying a car’s acceleration. But the pain and all of the rest of the mental lives of the players are in the patterns of their overt behavior (see Rachlin, 1985, for a behavioral view of pain, with commentaries and response). Some particularly wide patterns, such as consciousness, extend beyond the playing field but, in all cases, according to TEB, mental life is life—seen in an abstract, temporally, and socially extended way.

TEB is not a form of eliminative materialism. Although TEB would eliminate from scientific discourse the concept of consciousness as internal reflection, TEB also claims that the concept of consciousness (like all mentalistic concepts) is extremely useful and has an important place in scientific discourse. TEB claims that consciousness resides in abstract patterns of overt behavior. The difference between TEB and neural identity theories is not one of physical versus mental; we agree that the mental is real and is identical to an abstract pattern of activity of the organism. The difference is that, for modern monistic theories of mind such as neural identity theory, the pattern occurs in the brain, whereas for TEB, the pattern occurs over time in the organism’s overt behavior. It is not the word consciousness that TEB would eliminate from scientific discourse—still less from everyday speech. TEB does acknowledge the existence and reality of consciousness. Abstract entities, such as behavioral patterns, are at least as real as their components. It is not mental terms but rather introspection or internal reflection as a means of psychological investigation that TEB would eliminate. Nevertheless, reporting one’s mental states may be useful in everyday life. I may say, “I am angry,” or “I love you,” but not merely to report an internal state any more than I would say, “The grass is green,” or “The sky is blue,” merely to report an external state. Statements are made for reasons—practical or poetic; they have final causes. The reason, in the case of “I am angry,” and so on, is to predict one’s own future behavior (on the basis of one’s own current behavior and past behavior in similar circumstances). Such a prediction enables the hearer to react appropriately. It follows that a person (who is less observant of her own behavior than is someone close to her) may be wrong about an introspective statement. I might say, “I am angry,” and truly believe it, and my wife may say, “No you’re not,” and she may be right. It is introspection as a scientific method, not the reporting of one’s own mental activity that TEB rejects.

The neural identity theorist may claim that consciousness must be continuously present while we are conscious, whereas a behavioral pattern (such as playing in a Mozart quartet) involves pauses and stops. A violinist playing a Mozart quartet may pause occasionally but still she is playing the quartet, even when she is
pausing. A person who is conscious, on the other hand, seems to be continuously conscious. But neural firing, however rapid, is discontinuous. If consciousness were identical to that firing, it would also have to be discontinuous. If the neural identity theorist replies that the neural firing is so fast that consciousness bridges it over, she must have some nonneural conception of what that bridging over consists of—other than more, perhaps still more rapid, neural firing.

Consciousness is a vital property of our overt behavior with a vital function in our complex world. Our long-term patterns of behavior—sobriety, moderation, cooperation with others, morality, rationality, as well as the language that guides their organization—all evolved. (The power to learn, behavioral evolution, also evolved through biological evolution.) These patterns are what we would have to create in an automaton for it to leap over eons of biological evolution and be human. If such a machine were created in the future, the internal (or external) mechanism that will be the efficient cause of those patterns may very well turn out to resemble our actual nervous mechanism. Or it may not. But it will be the behavioral patterns of the machine, not its mechanisms, that count in determining whether it is conscious.

The view of consciousness proposed by Noë’s (2009) extended cognition theory extends (as does TEB) that the mind while expanding its reach, beyond the brain, into the peripheral nervous system and out into the world. But it does not consider a temporally extended view of cognition that extends behavior beyond the present moment into the past and future.

A kind of neural identity theory claims that consciousness arises only if a certain complexity of the nervous system is achieved. Consciousness, for this theory, need not occur in nerves themselves but could conceivably occur in a machine built of transistors and copper wires; what is necessary for consciousness is not the matter of brain physiology but the form of brain physiology, including its great complexity (Tononi & Edelman, 1998). For example, a computer might have an internal mechanism in which internal representations of input are placed on an internal “blackboard” or “global workspace” and scanned and operated on by an internal executive processor. According to Baars (1988), consciousness operates in (a more complex version of) this way. About this kind of theory, Dennett (2001) says, “Theorists must resist the temptation to see global accessibility as the cause of consciousness (as if consciousness were some other, further condition); rather it is consciousness” (p. 221). This is an identity theory where consciousness is held to be identical with some neural organization. Our brains might indeed work as Baars and Dennett suggest. But, again, if it were possible to construct such a (disembodied) mechanism in all its complexity, with no other possible movement, no function in the world, would consciousness occur in it? Psychologists and philosophers critical of neural identity theory do not believe it would.

Internal Events in Teleological Behaviorism

We frequently do talk or picture things to ourselves. Many mental acts, especially the “stream of consciousness,” seem to be nothing but internal monologues. But can talking or picturing things to oneself or covertly humming a tune bear the burden of accounting for all mental acts, including consciousness itself?
TELEOLOGICAL BEHAVIORISM

75

TEB does not object to inferring internal speech or internal picturing on the basis of observations of external behavior. TEB objects to labeling such inferred actions with a mentalistic vocabulary. A person talking to himself or herself is like a clock with the mechanism unconnected to the hands. From the viewpoint of TEB, that person is not thinking until the mechanism comes into contact with the environment and starts to serve its function. It is true that we first learn to read aloud and only then to ourselves. But, looking at a book and saying the words aloud or to oneself is reading only in the sense that a scanner “reads” a document. As Wittgenstein (1958) pointed out, real reading implies understanding. What is the difference between two people sitting quietly and reading Ulysses, one of them with a PhD in English and the other a high school dropout? In a narrow sense, these two would be behaving identically, yet they are (almost certainly) not thinking identically. That is, the patterns of their current and future overt behavior differ.

Conclusion

Psychology and philosophy should reject the covert view of the mind—what it means to hear a Mozart quartet, for example. At best, covert activities are only interpretable, not scientifically analyzable, in mental terms. One person’s good interpretation is another’s bad interpretation, and there is no one to decide between them. We may ask, in Descartes’s terminology, how can we obtain a “clear and distinct” view of our own minds? Such a view, I maintain, can be gotten not by looking inward with an MRI machine to some spot in our brains, but outward—to our overt behavior.


References


Received November 28, 2016
Revision received February 22, 2017
Accepted February 22, 2017

---

**E-Mail Notification of Your Latest Issue Online!**

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at https://my.apa.org/portal/alerts/ and you will be notified by e-mail when issues of interest to you become available!