

About Teleological Behaviorism

Howard Rachlin
Stony Brook University

Misconceptions abound about teleological behaviorism (TB). Because very few people other than the author publicly call themselves teleological behaviorists, the fault must be mine. The present article is an attempt to clear up those misconceptions. First I will try to indicate what teleological behaviorism is not. Then, in the form of six fables (loosely connected stories, allegories, analogies, fairy tales, and arguments), I will try to give the reader an understanding of what teleological behaviorism actually is.

Key words: behavioral evolution, mentalism, methodological behaviorism, operationism, S–O–R psychology, teleological behaviorism

Teleological behaviorism (TB) says that (a) an organism’s mental life resides in its overt behavior, in its temporal and social context; (b) behaviorism is the study of the overt behavior, over time, of the organism as a whole in its temporal and social context; and (c) behaviorists are the psychologists who are best equipped to study mental life.¹

I call myself a teleological behaviorist, and I view the argument above as a compound discriminative stimulus for my own behavior over at least the past 20 years. As my students will attest (and their word is crucial), it has been a highly valuable pattern to me, and I believe will be to anyone whose behavior becomes guided by it.

WHAT TELEOLOGICAL BEHAVIORISM IS NOT

Teleological Behaviorism Is Not a Form of Methodological Behaviorism

According to Graham (2010),

Methodological behaviorism is a normative theory about the scientific conduct of psychology. It claims that psychology should concern itself with the behavior of organisms (human and nonhuman animals). Psychology should not concern itself with mental states or events or with constructing internal informa-

tion processing accounts of behavior. According to methodological behaviorism, reference to mental states, such as an animal’s beliefs or desires, adds nothing to what psychology can and should understand about the sources of behavior. *Mental states are private entities which, given the necessary publicity of science, do not form proper objects of empirical study* [emphasis added].

First, TB is deeply concerned with mental *life* (if not mental *states*), as the reader will see when we come to the six fables later in this paper. Not only does TB concern itself with mental life but, according to TB, the mind plays a central role in a science of behavior. TB does not say that mental terms are necessarily “unscientific” or “opaque” or “vague.” Nor does it say that mental explanations are mere “interpretations” as opposed to scientific theories. Rather, mental life is central to a teleological view of behavior.

Second, TB does not view mental events as “private entities.” On the contrary, TB sees mental events as entirely public, as temporally extended patterns of overt behavior. TB does not see the mind as brain activity plus overt behavior, or as covert muscle twitches plus overt behavior. TB sees mental life as overt behavior patterns extended widely over time. Such temporally extended patterns, according to TB, are indeed proper objects of empirical study. TB accepts mental states as objects of scientific study but, once and for all,

¹ Much of the material for this article was borrowed from a book I am currently writing called *The Escape of the Mind* (Rachlin, in press).

Address correspondence to the author at howard.rachlin@stonybrook.edu.

rejects introspection as a path to scientific truth. From the viewpoint of TB, introspective reports are parts of patterns of overt behavior that are to be explained, like any other such patterns, in terms of contingencies of reinforcement. And this includes such apparent introspective certainties as “I know my mind is in my head,” “I am certain of my sensations,” “I know what I like,” and, “I think therefore I am.” Again, TB does not deny that we may know things, that we have sensations, or that we think. However, TB does deny that such events or states occur inside the organism and are available to that organism alone. TB positively asserts that those events occur in the patterns of the organism’s overt behavior and are available to anyone who can observe those patterns over extended periods of time.

According to TB, a person who is close to you and observes your overt behavior, including verbal behavior, over long periods of time (e.g., your spouse) may know what is on your mind as well as or better than you do. There is nothing *indirect* about those observations (cf. Moore, 2013). Thus, TB not only is unlike methodological behaviorism, it is just about as unlike methodological behaviorism as it can be. Contrary to methodological behaviorism, TB says that mental states are *not* private, and that they *do* form proper objects of empirical study.

Teleological Behaviorism Is Not a Form of Stimulus–Organism–Response (S–O–R) Psychology

The Dictionary of Psychology (Corsini, 2002) defines *S–O–R psychology* as follows:

A model of behavior. A stimulus touches the periphery of the body at a sensory organ (exteroceptive) or an internal stimulus occurs (interoceptive, proprioceptive), the organism routes the sensation through the nervous system (often including the brain) to the

muscles where some kind of overt or covert response occurs. (p. 947)

S–O–R psychology is most closely associated with the learning theorist E. C. Tolman. Skinner (1966), in the introduction to the seventh printing of his classic, *The Behavior of Organisms* (1938), had this to say about Tolman: “Although I continued to use the concept of drive for many years ... I regarded it simply as a convenient way of referring to *environmental* variables. Tolman, however, made it an *intervening* variable” (p. x). Between 1938 and 1966, Skinner abandoned the concept of drive altogether. However, for him, operant behavior was always a function of the *contingency* between response and reinforcer (the reinforcement schedule); for Skinner, a discriminative stimulus (S^D) did not cause the response or set off a chain of internal mediating events but signaled the presence or absence of one or another contingency. In all of the foregoing, TB agrees with Skinner and not with Tolman. Thus, TB is explicitly not an S–O–R psychology. And neither is TB an S–R psychology. The main difference between TB and Skinner is that TB takes a molar rather than a molecular approach to behavior. Whereas Skinner emphasized contiguity of response and reinforcer as the basis for behavior change, TB sees discriminative stimuli, responses (behavior), and reinforcer alike as temporally extended, observable patterns; behavior change, says TB, is due to correlations among these *patterns*, not contiguities among their *elements*.² The six fables in the latter part of this paper will illustrate the implications of this way of thinking

²Tolman called himself a “molar behaviorist.” In its molarity, TB is like Tolman’s behaviorism. But TB emphatically does not follow Tolman in his use of internal mediating events such as “drive” or internal representations (e.g., “cognitive maps”) as behavioral explanations.

for operant conditioning and show how mental life may be understood as operant behavior.

Teleological Behaviorism Is Not the Psychological Operationism of Boring and Stevens

Here is Stevens on operationism in psychology (Stevens, 1935):

Discrimination or differential response is the fundamental operation. ... By discrimination we mean the concrete differential reactions of the living organism to environmental states, either internal or external. Discrimination is, therefore, a "physical" process or series of natural events, and all knowledge is obtained, conveyed and verified by means of this process. (p. 518)

Stevens's operationism is a great improvement over the logical positivism that went before. In Stevens's operationism, nonphysical sensations were not the fundamental building blocks of all mental processes (as they were for Titchener); instead, physical discriminations were. This is a move toward behavioristic thought. But there are problems with Stevens's view. First, there is his apparent indifference to whether discriminations are internal or external. Internal observation is just another name for introspection. But rejection of introspection is (or should be) one principle that unites all behaviorists. Exactly what internal events are being observed? Are they retinal events? Are they brain states? Are they unobserved muscle twitches (covert acts)? And what internal homunculus is doing the observing? The concept of internal observation (introspection) is unbehavioral and, it seems to me, unoperational as well. And, I hope it is clear that TB involves no internal observation.

Second, Stevens seems to make a distinction in the above quote between discrimination and knowledge. If knowledge is not the same as discrimination (or a series of discriminations) but a state (presumably an

internal state) that is merely "obtained" by discrimination, then what is knowledge itself? Stevens does not answer this question.

Nevertheless, Stevens's operationism does contain a hint of TB's approach to mental terms. That hint is the indisputable fact that discrimination (of the external sort) takes time. A discrimination is not a single act; a discrimination is a temporally extended pattern of acts. If a driver stops only once at a red light, for example, he or she is not discriminating red from green lights (regardless of what may be going on inside of him or her). Any single instance of stopping could be due to numerous causes (e.g., stopping to get directions, to use a cell phone, because one foot has a cramp, or because the car ran out of gas). To discriminate red from green traffic lights, one does not just have to stop at a single red light; one has to do it many times. Moreover one also has to go at green lights, and one has to do that many times. The driver does not need to stop at all red lights and go at all green lights; there must merely be a statistically significant relation between red versus green and stopping versus going. In Skinnerian terms, red lights must be an S^D for one sort of act, and green lights must be an S^D for another sort of act. And, stopping and going are things that a whole person does; they cannot be brain states or covert responses. TB sees stopping at red and going at green as two intersecting behavior patterns.³ They must operate on the external environment; they must be operants. Why the external environment exclusively? Because that is where the reinforcement and punishment are. One will not get a ticket or be

³ Aristotle, in discussing discrimination of *white* from *sweet*, compares the overlap of the two patterns (discrimination of white from nonwhite and sweet from nonsweet) to a point of tangency between two circles. (See Rachlin, 1994, pp. 97–98, for a discussion of Aristotle's view of perception.)

TABLE 1

Comparison Between Radical and Teleological Behaviorism

| | Effective response– reinforcer relation | Mental terms | Inner causes |
|--------------------------|--|--------------|--------------|
| Radical behaviorism | Contiguity (molecular) | No | Yes |
| Teleological behaviorism | Correlation (molar) | Yes | No |

blindsided by a truck for covertly (but not overtly) running a red light. And, without reinforcement and punishment, one cannot be discriminating anything (cf. Gibson, 1979). Perception *is* that discrimination; that temporally extended pattern.

Teleological Behaviorism Is Not Mentalism

But let me back up. If by mentalism you mean any science that makes use of mental terms, then TB is mentalistic. But that is not what mentalism usually means. A good definition of *mentalism* in psychology is provided by Moore (2003):

In general terms, mentalism may be defined as an approach to the study of behavior which assumes that a mental or “inner” dimension exists that differs from a behavioral dimension. This dimension is ordinarily referred to in terms of its neural, psychic, spiritual, subjective, conceptual, or hypothetical properties. (p. 181)

It will be clear to the reader if it is not already that TB is not mentalistic by these criteria. Indeed, by its rejection of the oxymoronic concept of *covert behavior* (an imagined kind of movement that, by one part of its very name, is unobservable, a kind of movement that, by definition, is out of direct contact with reinforcing contingencies), TB is less mentalistic than those forms of behaviorism that resort to this concept.

Teleological Behaviorism Is Not Radical Behaviorism

To readers of *The Behavior Analyst*, this hardly needs to be said. Nevertheless, it may be helpful to specify the major lines of difference

as I see them. Table 1 summarizes these differences.

For Skinner’s radical behaviorism, the effective response–reinforcer relation is temporal contiguity. This perhaps reflects its origins in the study of the reflex. Skinner (1938) originally conceived of operant conditioning as another kind of reflex and as an extension of Pavlovian conditioning. On the other hand, TB owes its origins to experiments such as those of Herrnstein (1961), Herrnstein and Hineline (1966), and Baum and Rachlin (1969), in which programmed response–reinforcer correlations were shown to have direct effects on behavior (and which, I believe, are explained only with great difficulty in terms of contiguities).⁴ With regard to the admissibility of inner causes in a science of behavior, the explicit purpose of Skinner’s much-cited (1945) article, “The Operational Analysis of Psychological Terms,” was to interpret mental terms such as *thoughts* in operant-conditioning

⁴ Avoidance, for example, may be explained in terms of contiguities, as internal fear reduction by the avoidance response, the fear having been classically conditioned by the aversive stimulus. This explanation requires that the stimulus conditioned to elicit fear be reduced or eliminated by the avoidance response. Herrnstein and Hineline (1966) found reliable avoidance conditioning without any such stimulus. Explanation of this phenomenon in terms of contiguity must rely on hypothesizing an *internal* fear reduced by an *internal* removal of an *internal* conditioned fear stimulus. Such explanation sacrifices the principal advantage of behavior analysis: the observability of its variables. On the other hand, the negative contingency (in the wider temporal environment) between the rate of the avoidance response and the rate of the aversive stimulus is entirely observable over time.

language. Skinner was constrained in this effort by the necessity, imposed by his radical behaviorism, to find contiguous relations underlying all behavior change. Skinner did not look for causes in the wider temporal environment because he was constrained to explain all behavior in terms of response-reinforcer contiguity. But, as I have tried to illustrate, both above and below, with a wider temporal understanding, a molar response-reinforcer relation, in terms of correlations rather than contiguities, it is not necessary to interpret mental terms ("psychological terms") as unobserved events that occur within the organism. It does not matter whether those are inner spiritual, neurological, physiological, or muscular events. None of them are necessary to explain mental life as long as explanations are allowed on a molar level.

WHAT TELEOLOGICAL BEHAVIORISM IS

Let us turn now to the six fables. They are (a) the fable of Plato's cave; (b) the fable of Descartes, Jack and Jill, and the bad fairy; (c) the fable of Jeffery Grey and the behaviorist; (d) the fable of Sidney Morgenbesser, Dolly I, and Dolly II; (e) the fable that the New York Knicks will ever learn to play as a team; and (f) the fable of the time bug. Unlike Aesop, I will tell you the moral of each fable before I tell you the fable.

The Fable of Plato's Cave

Moral: Knowledge is overt behavior. Plato is sometimes said by philosophers to have invented the mind. You do not have to go this far to realize that Plato's thought about thought has had a great influence on how we think about thought. At one point, in his dialogue, *The Republic* (VII, 514a), Plato's famous "Allegory of the Cave" appears. The point of the allegory, according to Plato, is to

consider: "What is our nature in respect of education and its lack?" That is, learning. Plato is trying to answer the question: "What does it mean to learn something?"

In the allegory (Figure 1), a group of prisoners are chained to their seats in a cave. They are chained so tightly that they cannot move at all. (They cannot behave.) Moreover, they cannot see anything "of themselves or each other." All of the seats are facing one wall of the cave. The prisoners can see only this wall. Behind them is a platform on which slaves are walking back and forth. As they walk they are holding up chairs, tables, and all sorts of other objects. Behind the platform is a fire that projects shadows of the objects onto the wall.

Those were the days before movies, but I believe that Plato would not have objected to the image of the prisoners as the ultimate couch potatoes, condemned to watch movies and only movies for the whole of their lives. Then Plato imagines what it would be like for one of these prisoners to be freed from his chains and brought out of the dark cave and into the bright sunlight of our world. He describes the initial blindness and the subsequent enlightenment that the freed prisoner would experience. He interacts with real chairs. He scratches his nose, he engages in social and economic exchange with real people. That is, the world eventually functions for him as it does for us.

Then, in the allegory, the freed prisoner goes back into the cave where it takes him a while to readapt. He tries to explain to the unfreed prisoners what the shadows really represent. They cannot understand. They see him as a fool at best and, at worst, a grave danger. Of course, an allegory invites all sorts of interpretation. However, a critical point is that the prisoners in the cave do not see "anything of themselves or of one another" (*The Republic*, VII, 515a).

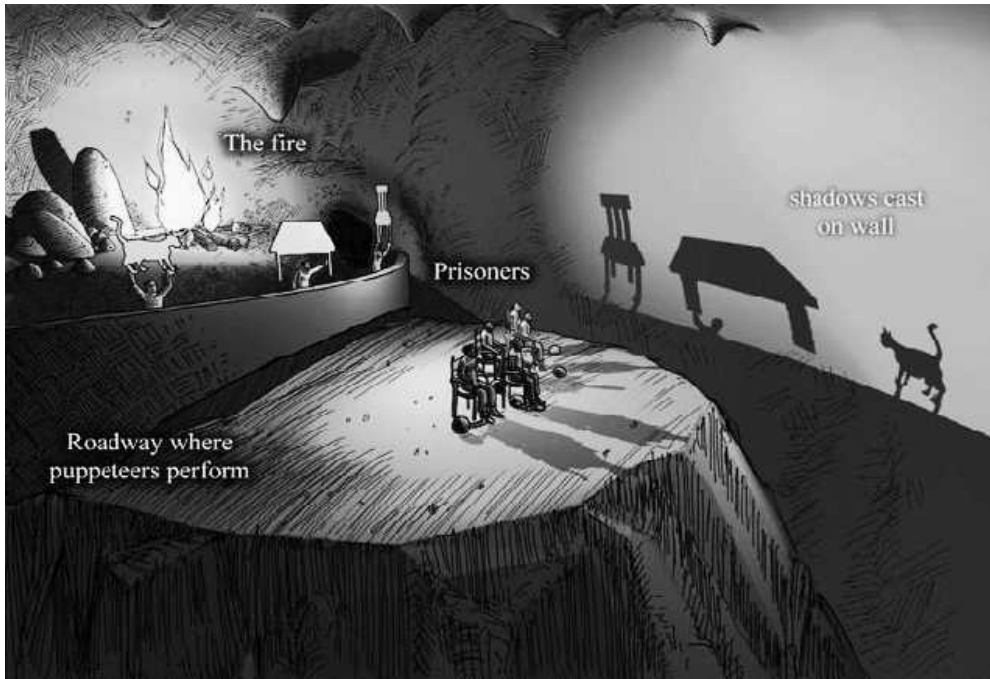


Figure 1. Plato's allegory of the cave. Drawing by Markus Maurer. Licensed under the Creative Commons Attribution-Share Alike 3.0 unported license.

The man who journeys out of the cave and becomes enlightened learns about his own reality at the same time that he learns about the reality of other people and the world. Plato is saying that all of us learn about ourselves in the same way that we learn about other people and about the world; by behavioral interaction, not by introspection.

The important point is not that the prisoners are seeing two-dimensional images and the freed man is seeing three-dimensional images. The allegory would apply with equal force if the prisoners were watching three-dimensional movies. The important point is that the prisoners are in chains, they cannot behave; that is, they cannot interact with themselves or one another. The objects that they see have no function in their lives. That lack of function is what makes those objects mere shadows of reality. For Plato,

function is the essential element of reality.

The allegory of the cave asks us to consider three worlds: (a) the world of the prisoners, (b) the world outside the cave (our own world), and (c) Plato's conception of an ideal world. By analogy, the world of the prisoners is to our world as our world is to the ideal world. We see the chair only as it exists in our world. But the "educated" person, who has expended the necessary effort, may come to interact with something approximating an *ideal* chair. According to Plato, you can know ideal objects, you can live in the ideal world, if (a) you are intelligent enough and (b) you expend the necessary effort. Moreover, according to Plato, we *should* live in that ideal world if we can. That world is not up in heaven, nor is it inside our heads. It is a mode of interaction with the actual environment in which we live. Plato is

saying that our behavior should conform to the abstract patterns of things and not to the particular elements. We will be happier if we do so, just as the freed prisoner is happier than the unfreed prisoners. In Plato's allegory, the journey of the freed man out of the cave is a metaphor for an intellectual journey, a teasing out from the environment of ideal forms defined by their function within their environment and community.

The point of the allegory, according to Plato, is to distinguish between education and its lack. As regards education, Plato objects to what, in modern terms, would be called the storehouse theory of knowledge. Education is "not like inserting vision into blind eyes," he says. It is not, in fact, an internal change at all. It is a change of behavior in the whole body: "The true analogy for this indwelling power of the soul and the instrument whereby each of us apprehends [learns] is that of an eye that could not be converted to the light from the darkness except by turning the whole body" (*The Republic*, VI, 518c). Or, as we might say, *the whole organism*. This passage is important because it is Plato's own explanation of the meaning of the allegory of the cave. A mental process (learning) corresponds to an interaction of *the whole body* with its environment rather than an interaction among parts of the body indicates that the freed prisoner learns because he now engages in a functional relation with the environment. Knowledge is an activity, not a state. The distinction between the prisoners and the freed man in the allegory is not just that one sees the world in two dimensions and the other in three dimensions, but that one is chained and completely passive and the other can interact with the environment including other people. Knowledge *is* that interaction. The mind that Plato invented is behavior in relation to the abstract patterns of

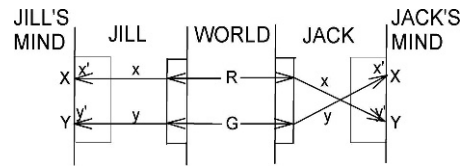


Figure 2. Jack and Jill both contemplating red or green lights. Jill is normal, but Jack's nerves are crossed so that red stimulates his brain in the place where green normally stimulates it and green stimulates his brain in the place where red normally stimulates it.

other people's behavior and of things in the world.⁵

The Fable of Jack and Jill and the Bad Fairy

Moral: The concept of mind enclosed as a prisoner within the body makes no sense. Once upon a time, in a far-away kingdom, the twins, Jack and Jill, were born. Jill was a perfectly normal baby, but Jack was not normal. (Assume for the time being that Jack's and Jill's respective minds exist wholly within their brains.) Just before Jack was born, the bad fairy (angry at not being invited to his bris) crossed up the connections between his retinas and his brain so that the nerves that transmitted a green signal went to the place in his brain where the nerves that transmitted a red signal would ordinarily have gone, and vice versa. Figure 2 diagrams this difference between Jack and Jill as they both contemplate red and green traffic lights.

When Jill sees a green object she feels the way everyone else does when

⁵ Think of Plato's allegory as the bull's eye of a target, then *The Republic* as a ring around it, then Plato's other works as a larger ring, then Plato's life as a still larger ring, and ancient Athenian society as a still larger ring. I am not a classical scholar, and I do not argue that the present interpretation of Plato as a sort of proto-behaviorist makes sense in the context of these rings (but see Rachlin, 1994, for a wider interpretation). Rather, I have plucked the allegory out of its natural context and placed it in the modern context of behavior analysis. There, *knowledge is behavior* makes an appropriate moral.

they see a green object (she has a green “quale,” as philosophers say) but when Jack sees the green light he feels the way everyone else does when they see a red light (he has a red quale). The internal mental state (if such a thing existed) of seeing red for Jack would be like that of seeing green for Jill, and vice versa. Jack exists in a world in which everyone else is constructed like Jill; Jack is the only one crossed up. The crucial question is this: Would Jack be hampered in any way by his unique physiology or would he get along perfectly well in the world? We may tend to think that he would at least be confused and have to learn to transform his red qualia into green behavior, and vice versa. But I think the reader will recognize that he would need to learn no more than Jill would.

Certainly, given the ultrasimplified physiology depicted in Figure 2, Jack would have no trouble communicating with Jill. Note that, because the brain exists wholly within the body, where light as such cannot penetrate, the brain activity X' has none of the properties of a red light and the brain activity Y' has none of the properties of a green light. The symbols X' and Y' merely stand for activity in different places in the brain. Unless God intervenes at some point, there can be nothing intrinsically reddish about R or greenish about G . The great philosopher Descartes realized this. Therefore, he concluded, God must intervene. For Descartes, all of the ideas we will ever have are innate within us. Stimulation from the outside merely wakes them up, as it were. If we assume Descartes was wrong about divine intervention in human consciousness, conscious states R and G can differ only in the sense that the letters X and Y differ, that is, as symbols that may mean anything you want them to mean. For Jill and the rest of us, R means red and G means green but, for Jack alone, R (the red quale)

means green and G (the green quale) means red. It would be no more difficult for Jack to learn to call R “green” than it would be for Jill to learn to call G “green.” For Jack, R would be a cool color, associated with grass, trees, mossy lakes, and so on; G would be a warm color, associated with the sun, fire, blood, and so on. In fact, as far as relating to the world goes, we don’t need the mind at all. Whatever the qualia R and G can do, the brain stimulation X' and Y' can do just as well. Not only is Jack at no disadvantage, but there is nothing short of brain surgery or MRI that will tell him he’s any different from Jill.

So the bad fairy has been foiled. Of course she is furious. With her evil power she rewinds time back to Jack’s birth. Again she crosses his nerves, but this time she crosses Jack’s motivational and emotional systems so that what causes Jill and the rest of us pleasure causes Jack pain (where pleasure and pain are conceived as taking place wholly within Jack’s body, out of direct contact with the world). And, what causes Jill and the rest of us pain causes Jack pleasure. Now poor Jack will suffer horrible pain whenever he eats an ice cream cone and will experience great pleasure when sticking a fork into an electric socket. I think everyone would agree that Jack would be truly screwed up if the bad fairy got her way. Figure 3 illustrates the situation.

But, just before Jack is actually born, the good fairy arrives at the bris. Although unable to directly undo the bad fairy’s physiological alteration, she adds her own manipulation. She arranges Jack’s nervous system so that he will be compelled to approach whatever gives him pain and avoid whatever gives him pleasure. Now, Jack will behave just like Jill. Both children will do their best to obtain and eat ice cream cones and avoid sticking forks into electric sockets. The difference will be that

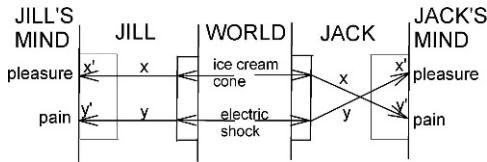


Figure 3. Jack and Jill both eating ice cream cones or subjected to electric shock. As in Figure 2, Jill is normal, but Jack's nerves are crossed. The nerves transmitting the taste and feel of the ice cream cone, which normally go to a place in the brain where pleasant stimulation goes, go to a place in Jack's brain where painful stimulation normally goes, and vice versa.

Jack will internally be in agony while smilingly eating the ice cream cone (he can't help it), whereas Jill will be internally experiencing pleasure while smilingly eating the ice cream cone.

Now, I have made up a lot of strange things in this fairy tale fable, but I think the reader will agree that this last scenario is truly weird. So we now have two weird notions: (a) Pain is overt behavior, and (b) a person may consistently appear to be experiencing pleasure while internally suffering extreme pain. Most people reject the first notion. The fable of Jack and Jill implies that if you reject the first notion you will have to accept the second. I personally find the first notion much easier to accept than the second. That is probably because I have been a behaviorist for most of my life. But I am also a pragmatist and I believe that the first notion is far more useful than the second for both treatment and understanding of pain (Rachlin, 1985).

You may ask, Cannot a person have a single episode of pain that is not only not reported to anyone but is not exhibited even as a brief exclamation, a grimace, a stumble, an intake of breath, and so on, that cannot be, even in principle, observable by another person? You are walking along the street, perhaps explaining something to a student who is attending closely to your every word and gesture. You suddenly have a shooting pain in your foot; can the

student notice absolutely nothing? I would answer this question in the negative. To me it is equivalent to saying that you may blink your eye internally without blinking your eye externally or jerk your knee internally but not externally. Of course it is possible for your nervous system to behave as it would if you were going to jerk your knee but then inhibit the overt motion at the very last synapse. But then you haven't jerked your knee. The reason why pain and pleasure are not essentially internal states at all but patterns of overt behavior. The first notion is only apparently weird, but the second is really weird. What lies behind this second notion? The false idea that the mind (whether a spiritual entity, or a structure in the central nervous system, or an event in the peripheral nervous system, i.e., a covert muscle twitch) is located somewhere within the body and is essentially private, not overt, behavior. Like a general, safe in his headquarters, sending and receiving messages to and from the front, the mind (isolated from the world, located deep in the brain, and served by millions of nerves) continuously receives messages from the sense organs and sends them to the muscles, but is never in direct contact with the world. This is the view of the mind that dominates current thought about thought that most people regard as settled. This is the view that creates the paradoxes and problems of mind and body. It is a thoroughly false view.

Let me repeat the moral of this fable: The concept of mind enclosed as a prisoner within the body makes no sense.

The Fable of Jeffery Gray and the Behaviorist

Moral: The notion that sensation, perception, and all mental activities are actually overt behavior patterns is not as absurd as it may seem. The

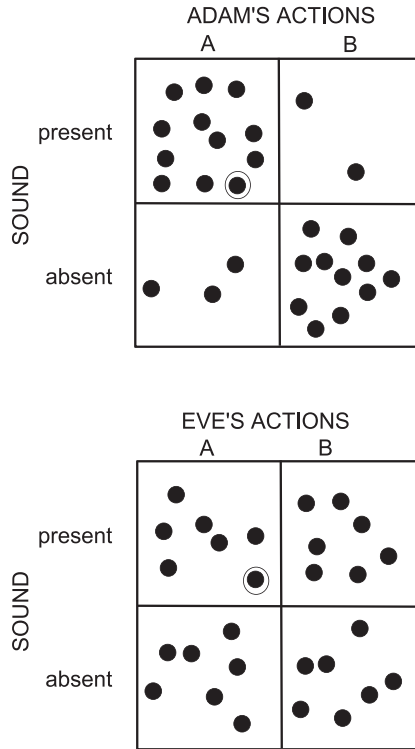


Figure 4. Adam has normal hearing and can discriminate sounds from their absence. Eve is deaf and cannot. The points represent occurrences of discriminative acts. At the moment, Adam and Eve are doing the same thing, as represented by the circled points. However, those identical acts are parts of two entirely different discriminative patterns.

British psychologist Jeffery Gray has claimed (as quoted by Staddon, 2001, p. 177) that once he asked a behaviorist the following question: “What is the difference between two awake individuals, one of them stone deaf, who are both sitting immobile in a room in which a record player is playing a Mozart string quartet?” Now it can be revealed that the behaviorist was me. I think that I mumbled something about the Mozart listeners’ future behavior. But here is what I should have said (the French call this a *pensée d’escalier*: a thought on the stairway of what you should have said at the time): I should have said, “The normal person (let us call him, Adam) is hearing the quartet and the deaf person (Eve) is not hearing it.” Surely Gray would have been very unsatisfied with this answer. He may well have gone on to

say, “But as a behaviorist you claim that hearing is behavior. How can one person be hearing something while the other not hearing it and yet they are behaving identically?”

Then what I should have done is to draw a diagram like Figure 4. To a behaviorist, what Gray’s question means is that you can parse at least some of Adam’s behavior into two categories: those that occur in the presence of sounds and those that occur in the absence of sounds. That is, for Adam, sounds can serve as discriminative stimuli. But for Eve, by assumption deaf, sounds cannot be discriminative stimuli. All of her actions are just as likely in the presence as in the absence of sounds. That is what it means for Adam to hear and for Eve to be deaf.

It is true that at the present moment Adam and Eve are doing

the same thing (as symbolized by the circled points). But these present acts are like two identical steps in two entirely different dances. The mentality of mental acts resides in an abstract pattern of overt behavior over time (the molar dance), and not in any specific, particular event (the molecular step). The Plato of the allegory of the cave would have said that Adam's and Eve's current identical postures are like projections on a wall, shadows, not real. He would have seen their differing discriminative patterns, however, as the place where their perceptions, or their failure to perceive, really lie.

*The Fable of Sidney Morgenbesser,
Dolly I, and Dolly II*

Moral: Our very consciousness, our sense of self, and our personhood are identical to our overt behavior in its social and temporal contexts. Once upon a time (about 20 years ago), I had occasion to deliver a lecture at the New School for Social Research in New York. In the lecture I argued that pain, like all mental events, is a pattern of overt behavior. I know that the idea of such an argument seems odd, not to mention counter-intuitive, but nevertheless that was what I was arguing, and still believe. But I will not defend the point here. (See Rachlin, 1985, for my argument, comments of various experts in the field, and my reply.)

In the audience was the philosopher Sidney Morgenbesser. Morgenbesser, now deceased, had published very little during his lifetime, but he was very, very smart. I can testify to that because I took several courses from him at the New School when I was a student there. One bit of evidence for Morgenbesser's cleverness is the following story: A philosopher giving a public lecture made the following statement: "In almost every language, two negatives make a positive, but in no language do two positives make a negative." Then a

voice from the audience called out: "Yeah, yeah." That was Morgenbesser.

So I was pretty nervous when he raised his hand after my talk. This is what he said: "Your argument is behavioristic, and behaviorism is refuted by the case of Dolly."

"What is the case of Dolly?" I innocently asked.

"Suppose," he said, "you were single and met a woman and went out with her. You had the greatest time of your life. She was wonderful in all ways. Beautiful, charming witty, laughed at your jokes, and so on. But then she reveals to you that she is a robot. Would you be disappointed?" My response to this was, "How far did we go?" But then I admitted that I would be disappointed. Who would want to marry a being without consciousness? Then we had some unenlightening discussion about what that meant. But what I should have said (this is another *pensée d'escalier*) is the following: "Imagine, Professor Morgenbesser, that you met a woman (Dolly II) and went out with her. You have the greatest time of your life. She is wonderful in all ways. Beautiful, charming witty, laughs at your jokes, and so on. But then she doesn't reveal that she's a robot. Instead you go out with her several more times. Then you ask her to marry you and she accepts. You have children (half human, half robot), grandchildren. She is the perfect wife in every way. Then on the morning of your 50th anniversary she takes sick and dies (rust has taken its toll). They do an autopsy and discover gears, pulleys, ratchets, a battery, and a nest of transistors inside of her instead of organic parts. Would you then regret then that you married her?" To me the answer is obvious: No, I would not regret marrying her. Assuming that Morgenbesser's answer were also no, this would mean that the cause of our mutual disappointment with Dolly I was not that her parts were inorganic rather than organic, but

rather our mutual implicit assumption that the patterns of her behavior would be constrained in time. That purely behavioral deficiency would have been her crucial departure from humanity. The reason why both Morgenbesser and I would have been disappointed with Dolly I is that current robotic technology has not even advanced as far as Dolly I. No one would be fooled by our best current version of Dolly I, even for a night. But let our technology advance to the point of creation of Dolly II and we might well be disappointed to discover that the girl (or boy) we met last night wasn't a robot.

The Fable That the New York Knicks Will Ever Learn to Play as a Team

*Moral: The New York Knicks will never learn to play as a team.*⁶ I have occasionally been asked, sometimes by members of ABAI, how patterns of behavior can be learned if they are not built up from chains of individual actions. In answer I would first like to point out that one person's individual action is another's temporally extended pattern of responses. Even a pigeon's key peck may be looked at as a complex pattern: orient to key, pull head back, open beak a little, accelerate towards key, decelerate as key is approached, and meanwhile attend to proprioceptive and visual feedback. If the complex pattern of key pecking can be reinforced as a whole, why shouldn't other more extended patterns (say a pigeon

playing ping pong) be reinforced (Locey & Rachlin, 2013)?

But let us consider a still more difficult question: How can a pattern be reinforced when every individual component of the pattern is dispreferred to its alternative? Such situations are common in cases of self-control. For instance, here is a quote from an autobiographical novel by Linda Yablonsky called *Junk*: "However good or bad you feel, heroin makes you feel better" (as quoted in the *New York Review of Books*, Sept. 25, 1997, p. 13). Yet over time Yablonsky prefers a pattern of abstinence from heroin to a pattern of heroin addiction. How is it possible to put together a pattern in which every component of that pattern is dispreferred to its alternative? How can people be self-controlled when each individual act of self-control is dispreferred to its (impulsive) alternative?

To answer that question, I would like you to consider an analogy between learning (which I here call *behavioral evolution*) and biological evolution. In biological evolution a similar problem arises in the evolution of altruism. Many animals behave altruistically. How can altruism evolve in animals when by definition every single altruistic act decreases the individual's fitness? One proposed answer to that question relies on a phenomenon called *group selection*. Imagine a population that consists of many individual tribes. Within each tribe, selfish people outcompete unselfish people. However, between tribes, those tribes whose members cooperate with each other unselfishly outcompete those tribes whose members are selfish. Or, to digress to another kind of selection, consider the National Basketball Association (NBA). All else being equal, teams that play unselfishly win more games, and therefore will survive in their league longer, than those teams on which each player plays selfishly. But, within each team, selfish players score the most points.

⁶ Written before the Knicks began the 2012–2013 season with a six-game winning streak. Their strategy seems to be approximately as follows: Carmelo Anthony (their star) plays on one side of the basket; the rest of the team plays on the other side. Anthony gets the ball if possible. If only one defender guards Anthony, he will probably score. If he is double or triple teamed, there must be an open man on the other side of the basket, and Anthony passes the ball off to him. This strategy takes great discipline on everyone's part and for now seems to be working.

So in biological evolution there may be a conflict between the survival of individuals and the survival of groups. Individual selection conflicts with group selection. Which kind of selection predominates depends on many factors, but one of the most important is rate of replacement. On NBA teams, the rate of replacement of players within teams is much greater than that of teams in the league. That is, players who play poorly are replaced on teams much more frequently than teams that play poorly are replaced in the league. So there is a force, at least, towards selfishness.⁷

In biological evolution there is a great dispute whether the conditions that favor group selection ever existed in human or nonhuman societies. I will not debate this question here (although I believe that such conditions have been and are present in human and nonhuman societies). But I believe that the conditions that favor group selection are present in behavioral evolution. In nature, individual responses are often tied together in bundles (e.g., the patterns of cumulative records Skinner found with various reinforcement schedules or so-called fixed action patterns of ethology) and can evolve as groups as a function of reinforcement. So even though every individual component of a pattern is not reinforced, or even if it is punished, the pattern as a whole may be reinforced. I believe that self-control can be learned in this way (not that we always learn it), and even altruism may be learned in the same way. We do not have to be born with innate altruistic tendencies. We can acquire them by group selection by reinforcement of cooperative patterns of behavior. And, I want to remind you, even a single key peck is already a complex pattern of behavior.

⁷This force is counteracted by the influence of coaches, managers, owners, fans, whose interests are usually the performance of the team as a whole.

The Fable of the Time Bug

Moral: The reason why we believe that consciousness is in our heads is that we believe (along with Descartes) that the cause of voluntary behavior is in our heads. Once we fully realize that the cause of voluntary behavior is in the social and temporally extended environment (as Skinner pointed out many times), we will recognize that consciousness is also there. Let me end this article with a science fiction fable. Imagine a planet on which all of the effort of an intelligent species was devoted to the breeding of plants and animals. And suppose that there existed on this planet an insect that was highly sensitive to both the time of day and the season of the year. In fact, the insect indicates the time by means of some externally observable signal; let us say that its color changes as the day progresses. The intelligent species has bred the insect to perfect the color change so that, although the insect might have originally needed to be recalibrated by exposure to a sunny day every once in a while, now (after many thousands of generations of breeding) it can live its life completely in the shade and still precisely indicate the time by its color. These insects would be very carefully preserved and fed. They might be taken aboard ships (along with a compass insect) to determine longitude and be carried around in a box that could be opened and consulted whenever the breeders needed to know the time. As the breeding of this insect progressed, the internal mechanism that caused its color to change would have evolved as well, becoming more and more precise. There would be a pacemaker, an accumulator (counting the number of pulses), a memory, a comparator, and so on. The bug could have such an internal timing mechanism, but it would be infinitely more accurate than ours.

The breeders would naturally use the insect's colors to indicate the time

just as we use the hands of a clock to indicate the time. They might know about the insect's internal mechanism or they might not. It would make no difference to their ability to breed the insect by killing off the relatively inaccurate ones and preserving and breeding the relatively accurate ones (they could check it with astronomical observations).

I claim that human consciousness evolved in a similar way, except by natural selection, imposed by the simultaneously evolving natural world and human society rather than by deliberate breeding. Just as time bugs were selected and preserved by their breeders, so conscious humans were selected and preserved in our world by their tribes. Just as the time appeared on the surface of the insects, so consciousness appears in our behavior. Certainly there is an internal mechanism that governs our consciousness; it might even exist in a specific brain location. And if it could be observed carefully, it might conceivably be "read" to predict overt behavior (as the wheels inside of a clock may be read to determine the time). But it would be a mistake to identify consciousness with such a mechanism. The intuition that consciousness is inside of us comes from the false belief that the origin of voluntary behavior is inside of us. As Skinner emphasized, the origin of voluntary behavior is actually in our environment; in our social system by which our behavioral patterns evolved. Consciousness, therefore, is in our overt behavioral patterns and not in the mechanism that generates those patterns. Consciousness can be directly known by observation of those patterns.

Skinner (1938) said,

As distinct from the other activities of the organism, the phenomena of behavior are held together by a common conspicuousness. Behavior is what an organism is *doing*—or more accurately what it is observed by another organism to be doing. ... Behavior is that part of the functioning of an organism which is

engaged in acting upon or having commerce with the outside world. (p. 6)

And, as Wittgenstein (1958) said, "If one sees the behavior of a living thing, one sees its soul" (p. 357).

REFERENCES

- Baum, W. M., & Rachlin, H. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, *12*, 861–874.
- Corsini, R. J. (2002). *Dictionary of psychology*. New York, NY: Brunner-Routledge.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Graham, G. (2010, Fall). Behaviorism. *The Stanford encyclopedia of philosophy*. Retrieved from <http://plato.stanford.edu/archives/fall2010/entries/behaviorism/>
- Herrnstein, R. J. (1961). Relative and absolute frequency of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267–272.
- Herrnstein, R. J., & Hines, P. N. (1966). Negative reinforcement as shock-frequency reduction. *Journal of the Experimental Analysis of Behavior*, *9*, 421–430.
- Locey, M. L., & Rachlin, H. (2013). Shaping behavioral patterns. *Journal of the Experimental Analysis of Behavior*, *99*, 245–259.
- Moore, J. (2003). Behavior analysis, mentalism, and the path to social justice. *The Behavior Analyst*, *26*, 181–193.
- Moore, J. (2013). Methodological behaviorism from the standpoint of a radical behaviorist. *The Behavior Analyst*, *36*, 197–208.
- Rachlin, H. (1985). Pain and behavior. *Behavioral and Brain Sciences*, *8*, 43–52.
- Rachlin, H. (1994). *Behavior and mind: The roots of modern psychology*. New York, NY: Oxford University Press.
- Rachlin, H. (in press). *The escape of the mind*. New York, NY: Oxford University Press.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century-Crofts.
- Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*, *52*, 268–277.
- Skinner, B. F. (1966). *The behavior of organisms* (Preface to the 7th printing). New York, NY: Appleton-Century-Crofts.
- Staddon, J. E. R. (2001). *The new behaviorism: Mind, mechanism, and society*. Philadelphia, PA: Psychology Press.
- Stevens, S. S. (1935). The operational definition of psychological concepts. *Psychological Review*, *42*, 517–527.
- Wittgenstein, L. (1958). *Philosophical investigations* (3rd ed.). G. E. M. Anscombe, Trans. New York, NY: Macmillan.