

## What Would It Be Like to Be IBM's Computer, Watson?

Henry D. Schlinger, Jr.  
California State University, Los Angeles

Rachlin (2012) makes two general assertions: (a) “To be human is to behave as humans behave, and to function in society as humans function,” and (b) “essential human attributes such as consciousness, the ability to love, to feel pain, to sense, to perceive, and to imagine may all be possessed by a computer” (p. 1). Although Rachlin’s article is an exercise in speculating about what would make us call a computer human, as he admits, it also allows us to contemplate the question of what makes us human. In what follows, I mostly tackle the second general assertion, although I briefly address the first one.

### TO BE OR NOT TO BE HUMAN

Without becoming ensnared in the ontological question of what it means to be human, let me just say that from a radical behavioral perspective, the issue should be phrased as “what variables control the response ‘human.’” This approach follows from Skinner’s (1945) statement of the radical behavioral position on the meaning of psychological terms. His position was that they have no meaning separate from the circumstances that cause someone to utter the word. Thus, when we ask what *perception*, *imagining*, *consciousness*, or *memory* is, we are really asking what variables evoke the terms at any given time. The same holds for the term *human*.

As one might guess, there are numerous variables in different combinations that probably evoke the

response “human” in different speakers and at different times. Rachlin claims that a “computer’s appearance, its ability to make specific movements, its possession of particular internal structures (e.g., whether those structures are organic or inorganic), and the presence of any nonmaterial ‘self,’ are all incidental to its humanity” (p. 1). However, it could be argued that one’s appearance and genetic (e.g., 46 chromosomes) and physiological structures (as behaviorists, we will omit the presence of any nonmaterial “self”) are not incidental to the extent that they, either alone or in some combination, evoke the response “human” in some speakers (e.g., geneticists, physiologists) at some times. For example, most people would probably call an individual with autism human because he or she has a human appearance and human genetic and physiological structures and behavior even though he or she may lack language and the consciousness that is derived from it. But because the variables that control Rachlin’s response “human” lie solely in the patterns of behavior of the individual organism or computer over time, we must wonder whether he would call this person “human.” If this conception of humanity is at all troubling, as behavior analysts, we would have to at least agree with him that “A behavioral conception of humanity is better than a spiritual or neurocognitive conception ... because it is potentially more useful” (p. 2). Once we accept his basic premise, we can move on to Rachlin’s second general assertion: that a computer may possess all of the attributes listed above.

---

Address correspondence to the author at the Department of Psychology, California State University, Los Angeles, 5151 State University Drive, Los Angeles, California 90032 (e-mail: hschlin@calstatela.edu).

### WHAT ATTRIBUTES ARE ESSENTIALLY HUMAN?

Rachlin includes sensing, perceiving, consciousness, imagining, feeling pain, and being able to love as "essential human attributes" (p. 1; later on, he includes memory and logic in the list), but what does he mean by "essential"? There are two possibilities. The first meaning, which I would call the strong view, is that only humans possess these attributes; that is, the terms are applied only to humans. Thus, as Rachlin argues, a computer that possesses them would be human. The second meaning (the weak view) is that although other animals (or computers) might possess some of these attributes, to be called human an individual must possess them all or must possess some (e.g., consciousness) that other organisms do not.

The term *strong view* (of essential human attributes) is meant to mirror, though not precisely, the term *strong AI* used by Searle (1980) to refer to one of two views of artificial intelligence (AI). According to Searle, in strong AI, "the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states" (1980, p. 417). (*Weak AI* refers to using computers as a tool to understand human cognition, and it is therefore synonymous with the information-processing model of modern cognitive psychology; Schlinger, 1992.) Rachlin appears to support the strong view of human attributes (that only humans possess them) when he writes, "These are human qualities by definition" (p. 10). Thus, if a computer possessed these attributes, it would be human. (Of course an alternate conception is that if a computer could possess them, then they are not essentially human.)

The strong view can be challenged simply by carrying out a functional

analysis of terms such as *sensing*, *perceiving*, *consciousness*, and so on, that is, determining the variables that control our typical use of the terms (corresponding to the terms' definitions), and then looking to see whether such variables occur in other organisms. Thus, without too much debate, I believe that we can at least eliminate sensation and perception from the strong view of essential human qualities. Sensation, as the transduction of environmental energy into nerve impulses, and perception, as behavior under stimulus control, are clearly present in most other species. The remainder of Rachlin's list of essential human attributes is trickier. Although some psychologists are willing to talk about animals being conscious and having feelings such as sadness and empathy, it is probably the case that such pronouncements are made based on some, but not all, of the behaviors exhibited by humans in similar situations. For example, regardless of any other behaviors, it is highly unlikely that other animals talk the way we humans do when we are described as being empathetic. Describing nonhumans with distinctly human-like terms is what leads some to level the charge of anthropomorphism.

Most behaviorists would probably not subscribe to the strong view of essential human attributes; that a computer possessing them would be human. The weak view, on the other hand, is more defensible, but not without its problems. Let us return to the question posed by Rachlin about whether a computer (IBM's Watson) can possess these attributes.

### COULD WATSON POSSESS ESSENTIAL HUMAN ATTRIBUTES?

In what follows, I address each of the essential human attributes that Rachlin believes Watson could possess, and argue that the responses

“sensation,” “perception,” “consciousness,” “feeling,” and “loving” as applied to Watson would be, at best, controlled by some, but not all, of the variables in humans that typically occasion the responses. Thus, the question of whether Watson can be made human may be moot. Perhaps a more relevant question is whether there is any justification for describing Watson, or any computer for that matter, with terms usually reserved for humans and some animals.

Before addressing the list of attributes, however, there is a more important issue to tackle, namely, the origin of almost all of the attributes listed by Rachlin. This issue is at the heart of Rachlin’s statement that “The place to start in making Watson human is not at appearance or movement but at human *function* in a human environment” (p. 4). This statement raises the question of just what a human function in a human environment is. With the exception of sensation, which is built in to biological organisms, the remaining attributes arise as a function of organisms interacting with their environment; and that environment consists of the natural environment as well as other organisms. In fact, one of the perennial problems in AI, at least for the first several decades of attempts to design and build computers that simulate human behavior, has been the failure on the part of researchers to recognize some important differences between computers and biological organisms, for example, that organisms have bodies that sense and act on the environment and their behavior is sensitive to that interaction; in other words, their behavior can be operantly conditioned (Schlinger, 1992). This is possible because organisms have bodies with needs and a drive to survive (Dreyfus, 1979). Of course, the “drive to survive” refers to the biological basis of staying alive long enough to increase the chances of passing on one’s genes. Based on

Dreyfus’s critique of AI, Watson would need something akin to this drive. Moreover, from a behavioral perspective, a computer would have to be constructed or programmed with unconditioned motivations and, like humans and other animals, be capable of acquiring conditioned motivations and reinforcers through learning.

Rachlin’s Watson II does have needs, but they are primarily to answer questions posed to him by humans. In addition, he needs “a steady supply of electric power with elaborate surge protection, periodic maintenance, a specific temperature range, protection from the elements, protection from damage or theft of its hardware and software” (p. 6). Getting needs met by acting on the world presupposes another significant human attribute, the ability to learn in the ways that humans learn. In other words, Watson II’s behavior must be adaptive in the sense that successful behaviors (in getting needs met) are selected at the expense of unsuccessful behaviors. Such operant learning is the basis of behavior that we refer to as purposeful, intentional (Skinner, 1974), and intelligent (Schlinger, 1992, 2003); as I have argued, any AI device (and I would include Watson) must be adaptive, which means “that a machine’s ‘behavior’ in a specific context must be sensitive to its own consequences” (Schlinger, 1992, pp. 129–130).

However, even if we grant Watson II’s ability to get these needs met through his “behavior” and his ability to “learn,” the question remains whether these attributes are functionally similar to those of biological organisms. Whatever the answer, the question is whether he would then be able to possess all the human attributes listed by Rachlin.

### *Sensation and Perception*

As mentioned previously, sensation refers to the transduction of environmental energy into nerve

impulses, and perception refers to behaviors under stimulus control (Schlinger, 2009a; Skinner, 1953). Following from these basic definitions, could Watson II sense and perceive? Obviously, unless Watson II was built dramatically differently than Watson, he would not sense in the same way organisms can. He would have no sense organs and sensory receptors that could respond to different forms of environmental energy. Using the information-processing analogy, we can say that information could be input (in auditory, visual, or even perhaps in tactile form if he were constructed as a robot). And although this type of input could become functionally related to Watson II's output, I do not think we would want to call it sensation. At best, it is analogous to sensation. On the other hand, if by "perception," all we mean is behavior under stimulus control, I think we could describe Watson II as engaging in perceptual behavior to the extent that his "behaviors" are brought under the control of whatever input he is capable of and assuming, even more important, that his behavior is sensitive to operant conditioning. However, even though describing Watson II as perceiving may be more accurate than describing him as sensing, it is still only analogous to what biological organisms do.

Speaking of Watson II behaving raises yet another problem. Would Watson II really be behaving? As the behavior of biological organisms is the result of underlying (physiological and musculoskeletal) structures, Watson II's behavior, like his sensing, would also be an analogy to the behavior of organisms. I do not think that mechanism is entirely unimportant, although I do agree with Rachlin that at least for most of our history it is not what has defined us as human, and that it might be possible to produce behavior with a different underlying mechanism that,

for all practical purposes, we would call human.

### *Imagining*

Even though constructing Watson II with attributes that resemble sensation and perception may be possible, arranging for the other attributes on Rachlin's list poses greater problems. First, however, let us agree with Rachlin by acknowledging that for the behaviorist, perception, imagining, consciousness, memory, and other so-called mental states or processes are really just words that are evoked by behaviors under certain circumstances (see also Schlinger, 2008, 2009a). As mentioned previously, to understand what we mean by these terms, we must look at the circumstances in which they are evoked.

With respect to imagining, the question is "What do we do when we are said to imagine and under what circumstances do we do it?" (Schlinger, 2009a, p. 80). The answer is that when we are said to imagine either visually or auditorily, we are engaging in perceptual behaviors that are evoked in the absence of actual sensory experience. Or, as Rachlin put it, "Imagination itself is behavior; that is, acting in the absence of some state of affairs as you would in its presence" (p. 7). For example, the behaviors involved in auditory imagining are most likely talking (or singing) to oneself, and in visual imagining, the behavior of "seeing" (Schlinger, 2009a). Note that the self-talk or seeing need not be covert (i.e., unobserved), but most often it is (Schlinger, 2009b).

Rachlin states that the behavior of imagining "has an important function in human life; that is, to make perception possible" and that "Pictures in our heads do not themselves have this function" (p. 7). But I would argue that he has it backwards: Perception (as behavior under stimulus control) makes imagining possible. In other words, we must

first act in the presence of certain stimulus events and have our behavior produce consequences before we can act in the absence of those events. We must first “see” a painting by Picasso before we can “see” the painting in its absence.

What would it take for Watson II to imagine? Simply speaking, he would have to be able to behave in the absence of stimuli. He would either have to “hear” (i.e., talk or sing to himself) in the absence of auditory stimuli or “see” in the absence of visual stimuli. In order to “hear,” he would need a verbal repertoire like that of humans, and to “see,” he would need some kind of visual system that would enable him to behave in the absence of the stimuli in ways that are similar to how he would behave in their presence. The verdict is still out as to whether this is possible.

### *Consciousness*

Let me start by saying that I agree with Rachlin that, “For a behaviorist, consciousness, like perception, attention, memory, and other mental activities, is itself *not* an internal event at all. It is a word we use” (p. 2). In fact, I said as much in an article titled, “Consciousness is nothing but a word” (Schlinger, 2008). (The rest of Rachlin’s statement—“to refer to the organization of long-term behavioral patterns as they are going on”—is open to debate, and I think we can address the problems raised by Rachlin without either accepting or rejecting his teleological behaviorism.) The critical point made by Rachlin is summed up in the following: “A computer, if it behaves like a conscious person, would be conscious” (p. 3). This statement evokes at least two questions: (a) What does a conscious person behave like? and (b) Could a computer behave like a conscious person? If the answer to the second question is yes, then we might further ask whether the

computer could behave like a conscious person without necessarily calling it “human.” A third possible question is whether a person can be a person without being conscious. Answering these questions requires some agreement about what it means to be conscious.

*What does a conscious person behave like?* In answering this question, radical behaviorists ask what variables cause us to say that a person (or any other organism for that matter) is conscious. In the previously referenced article (Schlinger, 2008), I listed at least three such situations. The first is when an organism is awake rather than asleep. This use of “conscious,” although not the most germane for our discussion, may still be applied, if only analogically, to Watson, just as it is to my Macintosh computer.

A second situation that evokes the response “conscious” is when an organism’s behavior is under appropriate stimulus control. For example, I say that my cat is conscious of his environment if he avoids walking into things, jumps on the bed, plays with a toy mouse, and so on. In this sense, animals are obviously conscious, and their behavior that leads us to say so has been operantly conditioned by interactions with the environment. (In this usage, the term is evoked by the same circumstances that the term *perceive* is. For example, saying “The cat perceives the mouse” is controlled by the same variables as saying “The cat is conscious of the mouse.”) Notice that the environment does not have to be a human environment; it can consist of the animal’s natural environment including other animals. Presumably a computer could be conscious in this sense as well if its behavior could come under the stimulus control of events in its environment as a result of interactions with that environment. For Watson II, its environment would presumably consist entirely of humans. It is this sense

of consciousness that interested Crick and Koch (2003) with their emphasis on visual perception.

A third circumstance that probably evokes the term *conscious* most often, and the one that is of most interest to consciousness scholars and laypeople alike, is the tendency to talk (i.e., describe) or imagine “to ourselves about both our external and internal environments, and our own public and private behavior” (Schlinger, 2008, p. 60). It is these behaviors that give rise to what consciousness scholars refer to as *qualia* or subjective experience and consist of what I believe a conscious person behaves like. That is, a conscious person is taught by his or her verbal community to answer questions about his or her own behavior, such as “What are you doing?” “Why did you do that?” and “What, or how, are you feeling?” (Schlinger, 2008; Skinner, 1957). As a result, we are constantly describing our behavior and private events both to others and to ourselves. Presumably, this is what Rachlin means by a human function in a human environment.

As Skinner (1945) first suggested, we learn to talk about private (i.e., unobserved) events in the same way we learn to talk about public (i.e., observed) events, that is, from others. In the case of private events, others have access only to the public events that accompany them. As a result, our descriptions come under the control, though not perfectly, of the private events. So, for example, we are taught to say “It hurts” when parents and others see either overt signs of injury, such as a cut or bruise, or when they observe us engaging in some kind of pain-related behavior, such as crying, moaning, wincing, and so on. Later on, we say “It hurts” only to the private painful stimulation. (Of course, it is also possible to say “It hurts” in the absence of any painful stimulation. Rachlin would still call this pain.) I believe that it is only because we

learned to say “Ouch” or “It hurts” from others that we actually are said to “feel” the pain, that is, the subjective experience of pain, as opposed to simply experiencing or reacting to the painful stimulation as my cat would. I think this is consistent with Rachlin’s statement, “To genuinely feel pain, Watson must interact with humans in a way similar to a person in pain” (p. 9). This sense of consciousness is simply an extension of perception, in that our verbal behavior is brought under the control of both public and private events dealing with ourselves.

Such self-talk is what I believe Descartes experienced that led him to state his famous *Cogito ergo sum* (I think, therefore I am) or, in behavioral terms, “I talk (to myself) about myself, therefore I am conscious of my existence.” Thus, although I might not agree with Rachlin about the details, I would agree with him that “consciousness is in the behavior, not the mechanism” (p. 3).

*Could a computer behave like a conscious person?* Based on the brief analysis presented above, for Watson II to behave like a conscious person, he would have to behave appropriately with respect to his entire environment, including the environment inside his skin. But therein lies the rub. We can grant that the computer should be able to describe its public behavior, whatever that behavior is, but what about private events? Without a sensory system that, in addition to exteroception, also includes interoception or proprioception, Watson II would not be able to describe private stimulation or, in other words, how he feels. And, unless he is constructed such that the mechanisms that produce behavior proximally (motor neurons, muscles) can function at reduced magnitudes without producing overt behavior, he would also not be capable of covert behavior and, thus, would not be able to learn to describe such behavior. So, at best, Watson II would

behave sort of like a human in that he could potentially be able to describe his overt behavior. But he would be handicapped in that he would have no private world to experience and, thus, to describe. But even if Watson II were able to describe only his overt behavior, would we call him human? As I suggested previously, I think that question is moot. It is probably best to skirt the ontological question and concentrate on whether Watson II could engage in human-like behaviors.

### WHAT WOULD IT BE LIKE TO BE IBM'S COMPUTER, WATSON?

In addressing this issue of qualia, Nagel (1974) asked, "What is it like to be a bat?" Based on the discussion above, the answer has to be "nothing." It is not like anything to be a bat, or any other animal, including preverbal or nonverbal humans, without self-descriptive behavior. As I have stated, "For the bat there will never be any qualia because there is no language to describe experience" (see Schlinger, 2008, p. 60). Even Dennett (2005) came around to this view of consciousness when he wrote, "acquiring a human language (an oral or sign language) is a necessary precondition for consciousness." Rachlin does not see it quite this way. According to him, we do not

know what it is like to be our brothers, sisters, mothers, fathers, any better than we know what it is like to be a bat ... if "what it is like" is thought to be some ineffable physical or nonphysical state of our nervous systems hidden forever from the observations of others. The correct answer to "What is it like to be a bat?" is "to behave, over an extended time period, ... as a bat behaves." The correct answer to "What is it like to be a human being?" is "to behave, over an extended time period, ... as a human being behaves." (p. 6)

Or, as Rachlin states elsewhere, "all mental states (including sensations, perceptions, beliefs, knowledge, even pain) are rather patterns of overt behavior" (pp. 3-4).

Although I understand Rachlin's point (after all, these are clear statements of his teleological behaviorism), I do not think such a position will be very palatable to traditional consciousness scholars. I believe that the position I have outlined here and elsewhere (Schlinger, 2008), although still perfectly behavioral, is closer to what consciousness scholars are getting at with their interest in qualia and subjective experience.

### CONCLUSION

Even though it may be possible to construct a Watson (Watson II) with attributes that resemble those in humans, the question of whether the resulting computer would be human is moot. A more practical question, as I have suggested, is whether there is any justification to describe Watson II with terms usually occasioned by the behavior of biological organisms, especially humans. But even then, the critical question is: What is to be gained by talking about computers using terms occasioned by humans? If we have to choose between the weak view of AI (that the main goal in building smart computers is to try to understand human cognition or behavior) and the strong view (that a computer with the essential human attributes mentioned by Rachlin would, for all practical purposes, be human), the weak view seems to be more productive. In other words, it would challenge us to analyze attributes such as perception, imagination, and consciousness into their behavioral atoms and the history of reinforcement necessary to produce them, and then try to build a computer (Watson II) that would interact with its environment such that those repertoires would be differentially selected. If we were successful, would we then call Watson II human? Rachlin's thesis is that we would. My point in this commentary is that such a conclusion is, at the

present time, too uncertain, and we would have to wait and see if Watson II would occasion the response “human.” I’m not so sure.

A more likely scenario, in my opinion, is that Watson II may be human-like in some very important ways. Regardless, the questions posed by Rachlin should help to pave the way for thinking about how to construct a computer that is most human-like. Rachlin is correct that, in order to do so, the computer must *function* like a human in a human environment. However, some of the so-called human functions mentioned by Rachlin (e.g., sensation and perception) are also possessed by other animals. And the functions he does mention that may be most distinctly human (e.g., consciousness) do not arise from interactions that differ in any fundamental way from those that are responsible for other behavior; in other words, the behaviors in question are selected by their consequences. Thus, the most important considerations in going forward with designing human-like computers is to build them with the ability for their “behavior” to be adaptive (Schlinger, 1992) and then see what happens.

## REFERENCES

- Crick, F., & Koch, C. A. (2003). A framework for consciousness. *Nature Neuroscience*, *6*, 119–126.
- Dennett, D. (2005). *Edge: The world question center*. Retrieved from [http://www.edge.org/q2005/q05\\_10.html#dennett24](http://www.edge.org/q2005/q05_10.html#dennett24)
- Dreyfus, H. L. (1979). *What computers can't do: The limits of artificial intelligence* (rev. ed.). New York: Harper Colophon.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, *4*, 435–450.
- Rachlin, H. (2012). Making IBM's computer, Watson, human. *The Behavior Analyst*, *35*, 1–16.
- Schlinger, H. D. (1992). Intelligence: Real or artificial? *The Analysis of Verbal Behavior*, *10*, 125–133.
- Schlinger, H. D. (2003). The myth of intelligence. *The Psychological Record*, *53*, 15–32.
- Schlinger, H. D. (2008). Consciousness is nothing but a word. *Skeptic*, *13*, 58–63.
- Schlinger, H. D. (2009a). Auditory imagining. *European Journal of Behavior Analysis*, *10*, 77–85.
- Schlinger, H. D. (2009b). Some clarifications on the role of inner speech in consciousness. *Consciousness and Cognition*, *18*, 530–531.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*, 417–457.
- Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*, *52*, 268–277.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.