ANTIREALIST ARGUMENTS IN BEHAVIOR ANALYSIS

François Tonneau
University of Guadalajara, Mexico

ABSTRACT: Some operant theorists have argued that the most fundamental concepts of behavior analysis have antirealist implications: for example, that stimuli have no physical properties, that we have no epistemic access to a physical world, that the world exists only in behavior, and that we are locked in our behavior. In this article, I show that such beliefs do not derive from behavior analysis. In particular, the concepts of stimulus and response employed in behavior analysis have no antirealist implications. Putative proofs to the contrary are seriously confused.

Key words: realism, stimulus, response, interdependence, relations, properties

In the last twenty years, some operant researchers who associate themselves with “pragmatism” and “contextualism” (e.g., Hayes, 1988) have promoted a brand of behavior analysis in which the traditional concepts of reality and truth are viewed with suspicion or rejected. Barnes and Roche (1994, 1997), in particular, have argued that the conceptual foundations of behavior analysis imply an antirealist view of the universe.

The sort of antirealism that Barnes and Roche want to establish concerns physical objects and the belief that such objects exist independently of us. Consider a wedge-shaped piece of metal, for example. A commonsensical assumption about this piece of metal is that it exists independently of any human being. Human beings were necessary to build the piece of metal, but once built it exists independently whether observed or not. Barnes and Roche (1997) emphatically reject this view (p. 545) and propose instead to characterize the piece of metal “only in terms of its behavioral functions” (p. 545), that is, in terms of an observer’s behavior with respect to it—such as looking at it and calling it “a wedge.” Going one step further, Barnes and Roche (1994) suggest that nothing ever exists independently of behavior: “the fundamental nature of the universe (or reality) exists as a behavioral event” (p. 167). Thus, what Barnes and Roche propose is a sort of behavioral solipsism in which one’s reality is contained in one’s behavior.

Traditionally, the arguments for antirealist views have drawn on a mix of philosophy, science, and the psychology of knowledge and perception (see the critical appraisals of Devitt, 1991, and Musgrave, 1989). Barnes and Roche (1994,
rather appeal to behavior analysis. They claim that their antirealist thesis is built into the very concepts of stimulus and response that are employed in behavior analysis (Skinner, 1935); according to Barnes and Roche, any behavior analyst who has ever used these concepts is thereby committed to antirealism. The argument runs roughly as follows. Echoing Dewey (1896), Barnes and Roche assert that “stimuli and responses are fundamentally inseparable” (1994, p. 167), in the sense that there cannot be a stimulus without a response (a “stimulus” without response wouldn’t be a stimulus). From this alleged “inseparability” (1997, p. 545), Barnes and Roche deduce that a stimulus is a purely functional state of affairs—a stimulus exists only to the extent that it is reacted to, and realism with respect to the physical world is false, or at least unwarranted: “If we talk of a real, physical universe, we are saying that stimuli have some form of existence beyond our behavior; this clearly contradicts behavior-analytic epistemology, in which there can be no stimuli (i.e., a physical universe) if there is no organism to provide responses that define those stimuli” (1994, p. 165).

It is safe to say that other operant researchers are reluctant to embrace such claims (e.g., Marr, 2003; Zuriff, 1995). In fact, Barnes and Roche (1994) admit that “most behavior analysts assume that there exists a real, physical, and ordered universe” (p. 165). Nonetheless, Barnes and Roche (1994, p. 165) maintain that their antirealism derives from widely accepted behavior-analytic assumptions about stimuli and responses. According to Barnes-Holmes (2000), for example, Skinner’s (1938) seminal treatment of the concepts of stimulus and response entails that “all events are defined or known as behavioral functions, instead of physical things that exist independently of behavior” (p. 197). If the latter assertion is correct, then there indeed is no room in behavior analysis for the realist hypothesis of a behavior-independent, physical world.

This, in essence, is Barnes and Roche’s brand of antirealism and their argument for it. Variants exist (such as the argument that because knowing is behaving, what one knows is always one’s behavior; see Zuriff, 1980), but they all point to a similar conclusion: behavior analysis is, or must be, antirealist. Here I will examine and refute the central argument of Barnes and Roche (1994, 1997) as well as its more popular variants. I will argue that the antirealist beliefs championed by these authors exemplify logical confusions and do not derive from behavior analysis (also see Hocutt, 1994). In particular, I will show that the concepts of stimulus and response used in behavior analysis do not have any antirealist implication.

**Relations and Properties**

As Phillips and Orton (1983) pointed out, claims about the “inseparability” of stimulus and response are reminiscent of late nineteenth century thought on “coordinations” (p. 161) and on a metaphysical thesis known as the *principle of internal relations* (p. 161). This thesis is difficult to state rigorously, but the general idea is that any relation is intrinsic to its terms. According to the principle of internal relations, for example, in any relation R between two entities A and B,
R is partly constitutive of the nature of A and B. Therefore neither A nor B can be “taken apart” from the relation, so to speak, because doing so would change their nature; removing the relation R between A and B would turn A and B into other entities.

The principle of internal relations turned out to be a cornerstone in the debates between idealists and realists of the early twentieth century (Spaulding, 1910a, 1910b). Believers in the principle of internal relations were led to the conclusion that the known was never independent of the knower, because what was known could not be the object as such but only the object-as-known (the cognitive relation, like all relations, being internal and therefore constitutive of its terms). Conversely, realists maintained that the cognitive relation was not internal but external; an object could enter the cognitive relation without being constituted by it (Baylis, 1929), and it could be known while remaining independent of the knower.

Although it would be possible to apply the principle of internal relations to behavior analysis with the antirealist implications that Barnes and Roche (1994) describe, behavior analysts have no reasons to do so. Not only is the principle of internal relations not part of behavior analysis, but the concepts of stimulus and response can be understood in a way that is fully compatible both with behavior analysis and commonsensical realism about physical objects. As far as the putative “inseparability” or “codefinition” of stimuli and responses is concerned, Barnes and Roche (1997) misunderstand its nature and import.

The issue can be clarified by first noting that any object has multiple properties and that some of these properties are relational—that is, they derive from the object’s sustaining particular relations with other entities. Being one mile north the Taj Mahal, for example, is a relational property (van Inwagen, 1993, pp. 33-35). From the fact that an object has multiple properties it does not follow that all of them are essential properties; some of the properties an object has can change without converting the object into something else. In particular, an object can lose or gain some of its relational properties without thereby losing or gaining all its properties. The persistence of the properties that survive the relational change, in conjunction with further conditions of spatio-temporal continuity, guarantees the continued existence of the object (Phillips & Orton, 1983).

To fix ideas, assume that John and Lisa are married. If they were to divorce, John would lose a relational property, that of being married. But John’s other properties would not necessarily change through the divorce; his blood type and date of birth, for example, would not change at all. Furthermore, the property of being married is not constitutive of John, since John can divorce while not ceasing to be John (although while ceasing to be married).

What properties are constitutive of John and guarantee his identity through time? Some philosophers maintain that physical constitution is essential, whereas

---

1 Alternatively, one could speak of intrinsic rather than constitutive properties (e.g., Hocutt, 1967). An intrinsic property of an object can, of course, involve relations among parts of this object. Thus, the property of being an atom of carbon involves relations among, and relational properties of, subatomic components (see Ellis, 2001).
others propose that John is extended through time and remains a single person as long as his temporal stages are connected by the appropriate causal relations (e.g., Loux, 1998). Either way, a person or object can lose some of its relational properties without becoming another person or object. John can divorce without ceasing to be John. Admittedly, for causal reasons the loss of a relational property, such as being married, can result in the loss of other properties, but whether and which of these properties are affected is an empirical issue (Baylis, 1929; Costello, 1911), not to be confused with the a priori principle of internal relations. Divorce can make John unhappy or even change his hair color, but it is still John who becomes single.

A Linguistic Confusion

Although the concepts of stimulus and response are widely acknowledged to be relational (e.g., Hocutt, 1967), relational properties have been a source of trouble in behavior analysis due to a linguistic ambiguity compounded with the careless use of terms such as “defining” or “define” (as when responses are said to “define” stimuli). The linguistic ambiguity is that we can refer to individuals (persons, objects) by mentioning some of their relational properties. Thus:

(1) My doctor told me not to smoke.

Clearly it is a person, not a relational property, who told me not to smoke; a relational property has no mouth and cannot speak. However, in order to refer to this person (say, Lorena), I mentioned her relational property of being my doctor—which is perfectly unambiguous because, as it stands, only one person has this relational property. As we have seen, the fact that Lorena is my doctor does not imply that Lorena has no other properties, and Lorena could cease being my doctor without ceasing to be Lorena.

The same point holds of pairs of individuals with their relational properties. When two persons entertain a relation R, they exhibit complementary relational properties derived from R. When a man and a woman marry, for example, they become husband and wife; that is, they acquire relational properties that they did not have before. By “husband” and “wife,” however, we may mean either the individuals or their relational properties, and failing to heed this distinction can lead to serious mistakes. Consider:

(2a) There cannot be a husband without a wife.
(2b) A husband does not exist independently of his wife.

I submit that (2a) is true; there cannot be a husband without a wife. Less ambiguously, a man cannot have the property of being a husband unless married to a woman who has the property of being his wife. Yet (2b) is false, since any husband exists independently of his wife. Lisa’s husband John, for example, exists independently of his wife, since he could divorce from her without ceasing to exist.
ANTIREALIST ARGUMENTS

How can (2b) be false if (2a) is true? Simply enough: in (2a) we are discussing the having of relational properties. Because these two properties, being a husband and being a wife, arise from a single relation (that of marriage), they always come in pairs; hence (2a) is true. In (2b), however, we are talking of the people who are husband and wife. These people can exist one without the other, as they did before getting married; hence (2b) is false.

In sum, it is important not to confuse relational properties with their bearers. From the fact that two relational properties come in pairs, it does not follow that their bearers must come in pairs. Nor does it follow that these bearers cannot be separated. Although being a husband and being a wife (the properties) always come together, husband and wife (the persons) can be separated; this happens every day, and this is called divorcing.

The fallacy of confusing relational properties with their bearers is committed by Barnes and Roche when they claim that stimuli and responses are “codefining” (1994, pp. 165-166, 168) and therefore “inseparable” (p. 167). If Barnes and Roche merely meant that an environmental event (say, E) cannot have the property of being a stimulus unless a corresponding part of behavior (say, B) has the property of being a response, then their thesis would be both correct and trivial, and no antirealist implications would follow. But Barnes and Roche (1994) also imply that the objects or events E and B cannot be separated, or do not exist independently of each other (p. 165), and here the argument is obviously fallacious. What Barnes and Roche are doing is moving illicitly from (3a) to (3b):

(3a) There cannot be a stimulus without a response.
(3b) A stimulus does not exist independently of responding.

The problem, of course, is that (3b) no more follows from (3a) than (2b) follows from (2a).²

Recall the claim of Barnes and Roche: “If we talk of a real, physical universe, we are saying that stimuli have some form of existence beyond our behavior; this clearly contradicts behavior-analytic epistemology, in which there can be no stimuli (i.e., a physical universe) if there is no organism to provide responses that define those stimuli” (1994, p. 165). One might just as well argue along these lines: “If we talk of real, physical husbands, we are saying that husbands have some form of existence beyond being married; this clearly contradicts the rule of marriage, in which there can be no husbands (i.e., physical men) if there is no ceremony to provide wives that define those husbands.”

² I first presented this argument in a talk at the Association for Behavior Analysis in 1998. Happily enough, nobody in the audience claimed that I was wrong about husbands and wives, yet I was told that the case was different with respect to stimuli and responses. I do not see why, and I know of no argument as to why the two cases should differ. Of course, the relations involved are different, being marriage in one case and stimulus control in the other (Hocutt, 1967). My argument applies to all cases of complementary relational properties regardless of the relation in which they are grounded.
In both cases, relational properties are confused with their bearers, or the fallacious inference is made that if an entity has some relational property then it does not have any physical property. Also notice the ambiguity of claiming that responses (or wives) “define” stimuli (husbands). Wives do not define husbands, and responses do not define anything. As Hocutt (1967) pointed out, one does not define a stimulus; rather, one defines the word, “stimulus,” or the concept of a stimulus.

Barnes and Roche’s antirealism carries the confusion one step further. Thus they write: “Does it really make any sense to talk of the fundamental nature of the universe as a non-behavioral event (i.e., as an independent reality)? It appears not, because as soon as you talk about the universe as non-behavioral event, it becomes a behavioral event. In other words, the universe can only ever exist in behavior” (1994, pp. 167-168). This claim is surprising. To say that an object becomes a stimulus is only to say that this object acquires a relational property that it did not have before. The relational property in question (e.g., being talked about) is behavioral, but this in no way implies that the object bearing the relational property is behavioral. Neither does it follow that this object cannot exist independently of its behavioral effects. That the properties of being-a-stimulus and being-a-response are relational has no antirealist implication at all.

Physical and Functional Characterizations

The co-occurrence of physical and functional (namely, behavioral) properties generally proves troublesome to Barnes and Roche. Thus they write, “in behavior analysis, all events are known or defined in terms of behavioral function, rather than [emphasis mine] as physical things that exist independently of behavior.” (1997, p. 545). This sentence sets up a false incompatibility between physical and functional characterizations. If Barnes and Roche mean that behavior analysts do not or cannot characterize stimulus instances in physical terms, then the claim being made is false. The instructions to authors of the Journal of the Experimental Analysis of Behavior, for example, recommend specifying stimuli in SI units (Système International d'Unités), a physical system if any (International System of Units, 2004). Notice that behavior analysts do not specify physical properties rather than, or instead of, functional properties. Behavior analysts specify both kinds of properties. They specify the constitutive properties of the objects or events being dealt with (otherwise nobody would know which objects or events are being described) and the relational properties of interest to behavior analysts, such as the property of being a discriminative stimulus (otherwise the resulting description would have nothing to do with behavior analysis).

My claim that behavior analysts describe the constitutive properties of stimuli along with the relational ones is not an expression of personal taste but a factual statement. No behavior analyst has ever published a report in which the only property of the stimulus object ever mentioned is that of being a stimulus for a response, or in which the only description made of the experiment is that the stimulus stimulated and the response was the response. Clearly, constitutive
properties (what stimulus are we talking about? what response?) must be mentioned along with relational ones; otherwise the resulting science would be vacuous (Gibson, 1967).

Barnes and Roche’s (1997) further claims demonstrate the hopelessness of pretending that stimuli have no constitutive properties. After stating that in behavior analysis a wedge “is defined only in terms of its behavioral functions” (1997, p. 545), Barnes and Roche suggest that the wedge “may be defined as a discriminative stimulus for a particular response, such as picking it up.” But since the “it” in this sentence obviously refers to the wedge, the latter (as a physical object) has not been eliminated from Barnes and Roche’s description. Their next proposal, namely that the wedge “may be defined as a reinforcing stimulus for other responses, such as pointing at the wedge” (p. 545) similarly fails, since the description of the response (“pointing at the wedge”) again mentions the wedge. In order to carry out their antirealist program, even in such a simple example, Barnes and Roche would need to characterize the relevant behaviors without ever using terms like “it” or “the wedge.” Unsurprisingly from a realist standpoint, Barnes and Roche (1997) have not done so.

Can Two Persons Know the Same Thing?

Barnes and Roche’s argument is reminiscent of earlier philosophical debates in more than one way. In an article supportive of realism, Russell (1911) wrote of the proponents of internal relations:

They say that if A is the father of B and the son of C, it is not strictly the same entity which is father and son, but that it is “A quâ father of B” who is the father of B, and “A quâ son of C” who is the son of C. This doctrine also is denied by those who advocate external relations. They would argue that “A quâ father of B” is a complex containing the constituent A, and “A quâ son of C” is also a complex containing the constituent A. Thus the attempt to avoid an identical constituent in two complexes break down. (pp. 159-160)

The issue of whether the same object can serve as a stimulus for different observers reappears in Barnes and Roche’s (1997) analysis. At one stage of their antirealist argument (pp. 545-546) they consider the case of two drivers reacting to what seemingly is the same red light. On the ground that the situation involves two distinct discriminative functions (one for each driver), Barnes and Roche infer that “in a sense, there are two red lights—one in each behavioral stream” (1997, p. 546).

Apparently, the claim is made that the same physical object cannot share discriminative functions with respect to different behaviors (also see Barnes-Holmes, 2003). But why not? Notice that Barnes and Roche (1997) do not derive their claim from empirical properties of the stimulus situation (e.g., the fact that the two drivers interact with the red light from different positions and are therefore stimulated by different light patterns). Rather, Barnes and Roche argue on logical
grounds—on the basis of alleged logical properties of the concepts of stimulus and response.

Barnes and Roche’s antirealist thesis again confuses relational with constitutive properties. Claiming on logical grounds that the same physical object cannot be a stimulus for different persons is like claiming on logical grounds that the same physical man cannot marry two different wives—but of course he can, however illegal; this is called bigamy. If John marries Anna and Edna, John becomes Anna’s husband and Edna’s husband. We have two relational properties, one physical man. There is no sense in which there are two husbands (i.e., two physical men, one in each marriage), although, of course, there are two relational properties (being Anna’s husband and being Edna’s husband, respectively).

Notice that the doctrine of internal relations would make bigamy impossible, since in the very act of marrying a different wife a man would thereby become a different man! Undoubtedly the man can “become another husband,” but only in the sense of acquiring a relational property that he did not have before. The man himself exists independently of either wife—or so I hope.

The Prisoner Variant

An equally flawed argument in favor of antirealism involves the assumption, uncontroversial among behavior analysts, that knowing is a behavioral phenomenon. On the ground that “to know is to behave in a certain way” (Zuriff, 1980, p. 342), it is inferred that “human knowledge of the world consists of responses to that world” (p. 342) and that human beings are somehow locked in their responding. Because “humans cannot transcend their own behavior to step out of the causal stream” (Zuriff, 1980, p. 342), they have no epistemic access to a behavior-independent world; in Barnes-Holmes’ words, “not even the behavioral scientist can escape his or her behavioral stream and make direct nonbehavioral contact with an ontological reality” (2003, p. 148). Thus the argument leads to the same sort of antirealist conclusions that Barnes and Roche (1994, 1997) want to establish—that the world exists only in our behavior, we have no access to a response-independent reality, etc.

Here the fallacy differs slightly from the ones we have reviewed. The premise of the argument is correct, at least from a behaviorist standpoint. Assuming (as behaviorists must assume) that knowing is a behavioral phenomenon, it is indeed true that in order to know an environmental object or property (E) I must react to it with some response (B). But from this it does not follow that I know B instead of E; my knowledge of E might be identical with E, as the theorists of direct perception argue (e.g., Tonneau, 2004). From a direct realist standpoint I cannot know E unless I react to it with some response B, but it is still E, and not B, that I know. I know E because of my response to it, not in spite of my response to it; far from being an obstacle to knowledge, my response to E is a necessary condition for knowing E itself.

Take the case of eating an apple. Certainly I need my teeth in order to eat an apple. Yet it would be absurd to argue that because I eat the apple with my teeth,
therefore I eat my teeth instead of the apple. It would be even more absurd to claim that in order to eat apples I would need to get rid of my teeth, and that since this is generally impossible, I cannot eat apples. The fact is that whenever I eat an apple with my teeth, I do not eat my teeth. I do not even eat the apple through my teeth. Rather, I eat the apple with my teeth, and it is the apple that I eat, not my teeth.

So the fact that it is behavior that allows us to know an object has no antirealist implications. We may know an object E by reacting to E with B, thereby knowing E itself without being locked in B. A behavioral conception of knowing (“I know E to the extent that I react to it”) is perfectly compatible with direct realism about E, and the world in general. True, in order to know my environment I must behave, but this does not imply that I know my behavior instead of the environment, still less that I am a prisoner of my behavior.

Antirealists may, of course, postulate that when trying to know an object E with a response B, it is B instead of E that we know. Unsurprisingly, antirealist assumptions lead to antirealist conclusions, but the latter do not follow from a behavioral analysis of knowledge per se, which is neutral with respect to the realism controversy. Again, the antirealist conclusion derives from unarticulated premises or confusions instead of behavior-analytic assumptions.

**Conclusion**

Inferring that stimuli have no constitutive properties on the ground that they have relational ones (Barnes & Roche, 1994, 1997) would leave behavior analysis in a sorry state as a science. As I pointed out, behavior analysts always specify the constitutive properties of stimuli along with their behavioral ones. Behavior analysts do not have much of a choice. In an exchange with Hocutt (1967) that was also a prescient commentary on the concept of stimulus, Gibson (1967) wrote:

An object, an environmental source of possible stimulation, cannot be thought of as simply one member of a reciprocal pair of terms each implying the other. That would be circular. It means conceiving the perceptual response as dependent on the environment while conceiving the environment as dependent on the perceptual response. It suggests the doctrine that the essence of an environmental object is to-be-responded-to; it suggests, with Bishop Berkeley, that its *esse* is *percipi*. Has not Hocutt noticed how many stimulus-response psychologists fall straight into the arms of Berkeley when they begin to theorize? (p. 533)

Gibson’s warning applies well to the antirealist brand of behavior analysis. Properly understood, the concepts of stimulus and response are relational but have no antirealist implications; speculating on implications that don’t hold can only hamper the philosophical clarification of behaviorism.

Now, I have not been trying to prove that the sort of antirealism that Barnes and Roche (1994, 1997) want to establish is false. Maybe the world exists only in behavior, maybe we cannot have access to the world independently of our behavior, maybe we are locked in our behavior stream, etc. However, these
conclusions do not follow from the logical properties of the concepts of stimulus and response, they do not follow from any coherent practice or theory in behavior analysis, and they do not follow from the assumption that knowing is a behavioral phenomenon.

The arguments of some behavior analysts in favor of antirealism exemplify the doctrine of internal relations or confusions between relational and constitutive properties. Whatever the reasons some behavior analysis may have for becoming antirealists, these reasons do not arise from behavior analysis.

References

ANTIREALIST ARGUMENTS


