BOOK CHAPTER - FINAL DRAFT

To appear in:

E. P. Charles (Ed.) (2011). *A new look at New Realism: The psychology and philosophy of E. B. Holt*. Piscataway, NJ: Transaction Publishers.

Author's address:

Dr. François Tonneau
Universidade do Minho
Escola de Psicologia
Campus de Gualtar
4710-057 Braga
PORTUGAL

E-mail: ftonneau@psi.uminho.pt

Holt's Realism: New Reasons for Behavior Analysis

François Tonneau
Universidade do Minho

> A consciousness is the group of (neutral) entities to which a nervous system, both at one moment and in the course of its life history, responds with a specific response. (Holt, 1912, p. 373 in *The New Realism: Coöperative Studies in Philosophy*.)

Behavior analysis must be supported by a realist philosophy, or so I argue. The New Realism that Holt and his colleagues launched in 1910 ("new" because it succeeded a long idealist eclipse) enjoyed only a few years of existence, and behaviorism in any radical form has been claimed to be dead for a while; defending neorealism as the proper philosophical foundation of behavior analysis may thus seem pointless. Two arguments nevertheless suggest that my project is not entirely in vain. First, strains of Holt's neorealism remain present in James Gibson's ecological psychology, and the death of behaviorism has been announced too many times to be trustworthy (Roediger, 2004). However marginal with respect to mainstream cognitive science, ecological psychology and behavior analysis show little evidence of extinction, and within a limited niche their accomplishments are genuine enough to warrant serious consideration. Secondly, in spite of important differences, the

emerging concepts of extended mind (e.g., Clark & Chalmers, 1998) and enactive cognition (e.g., O'Regan & Noë, 2001) show that the grip of representationalism on cognitive science is weakening (Menary, 2006). The peak of the physical-symbol system hypothesis (Newell, 1980) has definitely passed, and the role of the environment in the demarcation of mental processes is discussed with increasing urgency (Chemero, 2009). The nature of the current debates in cognitive science thus points toward a reconsideration of the concepts of direct perception, memory, and consciousness in relation to behavior.

Can Holt's realism prove of value to behavior analysis? This question can be addressed only through a proper understanding of the former and the latter, separating the philosophical core from historical accidents. Many of the contemporary features of behavior analysis, for example, were acquired through the contributions of individual scientists and reflect their own judgments and misjudgments. Thus, behavior analysts' views on the importance of reinforcement can and do vary (e.g., Malone, 1978), and one can propose alternatives to Skinnerian concepts while remaining a behavior analyst (e.g., Rachlin, 1973). There is one feature of behavior analysis, however, that seems indispensable. No researcher could stop studying relations between environment and behavior and still claim to be doing behavior analysis. Behavior analysis places no restrictions on how these relations can be studied or described, but it does prescribe that these relations are studied. Thus, behavior analysis embodies a mode of explanation that appeals fundamentally to the environment, motivating a description of the field as an environment-based psychological theory (Hineline, 1990). Environment-based psychological theory assumes that mental phenomena (the phenomena that are commonly deemed "psychological") consist of relations between environment and behavior. Notice that it is not merely assumed that the relations between environment and behavior are somehow relevant or important to psychological issues; rather, as in neorealism and ecological psychology, relations between environment and behavior are claimed to be *constitutive* of psychological issues (Smith, 1988).

The environment-based thesis opposes any framework in which the mind mediates causally, instead of being constituted by, relations of environment to behavior. To Anderson's (1991) claim that "there is a mind between the environment and behavior" (p. 513), for example, the behavior analyst is likely to object that the mind *consists* of the relation between environment and behavior (e.g., Rachlin, 1991, 1994). It is not surprising, therefore, that some of the strongest criticisms of behavior analysis have arisen from proponents of computational and representational approaches to cognition, starting with Chomsky's (1959) review of *Verbal Behavior* (Skinner, 1957). Many of the criticisms focus on what Pylyshyn (1984, 1987) has called the stimulus independence of behavior, an elaboration on Chomsky's argument (originally made with respect to linguistic utterances) that the overt relations between environment and behavior are too unstable or disordered to make for meaningful explanations.

Excluding the environment from the constituents of psychological phenomena, however, is both older and more pervasive than current representational frameworks. It starts with the psychology of perception to reach epistemology and philosophy at their core. Psychophysical scaling, for example, encourages the belief that the environment in which we live is not an object of knowledge. In perceptual experience we are supposed to be aware not of external objects but of subjective magnitudes or internal qualities brought to

us by nerve bundles. After the comparatively benign distortions of sensory scaling come the surprises of optical illusions and the full-fledged puzzles of hallucination, phantom limbs, imagery, and dreams.

All forms of apparent disconnection from the environment, from benign to severe, encourage the epistemology of indirect realism, according to which we are never aware of *anything* except "representations," conceived either as immaterial sense data or as neural signals. Indirect realism then leads to idealism or philosophical skepticism about the external world, as in Descartes' *Méditations Métaphysiques* (in which he ponders the possibility that the experienced world is only a dream). If the Cartesian dialectic is correct, then there is little hope for behavior analysis to succeed. Not only is it impossible, short of tenuous inference, to study relationships between environment and behavior, but this shared environment can hardly be known and its very existence is in doubt (Katz & Frost, 1979).

At this stage of the argument, behavior analysts may be tempted to dismiss the entire debate as metaphysical and fruitless, but they ignore metaphysical issues at their own peril (Tonneau, 2005a). The implications of indirect realism will not disappear simply by refusing to examine them. If indirect realism, idealism, or skepticism about the external world are correct, then behavior analysis is at best severely limited and at worst impossible. The practice of behavior analysis may remain what it is, but will remain so only at the cost of theoretical or philosophical incoherence. The only alternative to complacency is to focus the issue sharply and return to realism as a guiding framework for behavior analysis—but this realism must be of the direct, not indirect, kind (Katz & Frost, 1979).

Behavior Analysis and Perception

However unpalatable in its implications, indirect realism is buttressed by an impressive array of arguments. Most of them revolve around the issue of nonveridical perception, from subjective sensations to dreaming. Given this background and the importance of the challenge, one would expect behavior analysis to have taken a proactive stance with respect to the issue of nonveridical perception. A successful rebuttal to the argument from nonveridical perception should presumably tackle psychophysics and illusions first, and then extend to mental imagery, hallucinations, and dreams. In all cases, the behavior analyst's aim in the debate would be to show, though a combination of empirical research and theory, that the phenomena that seem incompatible with direct realism are actually consistent with it (and therefore behavior analysis as defined above). The defense, if successful, would require detailed attention to epistemology and metaphysics (e.g., Turvey, Shaw, Reed, & Mace, 1981), and would in turn strengthen both direct realism and the further study of relations between environment and behavior.

This has barely happened. Here I cannot discuss all of the underlying reasons (they are complex and many), but I can sketch how the issue of nonveridical perception stands in behavior analysis and discuss a few of the associated problems. A first problem is that in contrast to ecological psychologists (e.g., Gibson, 1960, 1979), behavior analysts are generally disinterested in stimulus structure or in specifying the exact nature of the operative stimulus. Following Skinner (1938), behavior-analytic research on stimuli usually evaluates how they come to affect responding through operant or Pavlovian processes (Honig & Urcuioli, 1981). The topic of generalization gradients (in which a

response is reinforced in the presence of a target stimulus and then tested against a range of stimulus variants) actually decreased in prevalence in the last two decades, to be replaced by a focus on how stimuli acquire new functions through participating in complex choice relations (e.g., Sidman, 2000). Little of this involves perceptual research about the nature of the effective stimulus.

Undoubtedly operant methods have proved useful for the study of animal perception, but when psychophysical analyses are attempted in behavior analysis they often involve signal-detection theory in its traditional form (Blough & Blough, 1977). Consistent with indirect realism, signal detection theory assumes that external stimuli are converted into internal sensations characterized in terms of a probability distribution along a decision axis. It is these internal sensations that are supposed to guide the animal's choices.[1] Aside from one attempt by Zuriff (1972), which focused on philosophy and did not inspire any concrete research program, I am unaware of any behavior-analytic treatment of psychophysical scaling.

The situation with respect to more extreme forms of nonveridical perception is similar. The *Journal of the Experimental Analysis of Behavior* has published only two empirical articles related to imagery, one dealing with afterimages in the pigeon (Williams, 1974) and the other with reaction times during mental rotation (Cohen & Blair, 1998). Research on dreams is nonexistent. To this paucity of empirical research we must add the inadequacy of Skinner's conceptual stance toward the issue of perception. Instead of developing an approach to nonveridical perception that would be consistent with the remainder of behavior analysis, Skinner (1974, 1989) instead leant toward indirect realism. He assumed that in mental imagery and dreams, for example, we were aware not of any external object but instead of our own "sensing," conceived as an internal physiological activity (e.g., Skinner, 1989). As Hayes (1994) pointed out, Skinner's insistence that the study of perception (or "sensing") should be left to brain science (Skinner, 1989, p. 14) effectively removed the issue from the domain of behavior analysis.

Skinner's forays into indirect realism are disappointing, because his own work on behavioral causation, which stresses the importance of historical factors in the explanation of behavior (e.g., Skinner, 1972), leads to an attractive alternative to the internalist explanation of nonveridical conscious contents. Although much remains to be done to develop a full-fledged direct realist theory of consciousness within the boundaries of behavior analysis, its seeds can be found in some contemporary approaches to operant performance. A good example of it can be found in White's (1985, 1991) experimental work on short-term memory. Animal memory is often studied with delayed match-to-sample tasks in which a stimulus is shown to the animal and then removed. After a few seconds, the animal is given the opportunity to choose among various options; the nature of

[1] Interestingly, some of the predictions of signal detection theory can be emulated by generalization models with a parameter expressing the extent to which the behavioral effects of a stimulus propagate to others (Davison & Tustin, 1978; Nevin, 1981). The value of the generalization parameter in these models, however, is not derived from more basic assumptions about the nature of effective stimuli, and more recent behavior-analytic work reverts to the assumption that stimulus generalization arises from distance in a psychometric space (Nevin, Davison, & Shahan, 2005) or is guided by internal sensory effects (White & Wixted, 1999).

the correct response depends on the identity of the stimulus previously shown. Provided that the delay between the stimulus and the opportunity to respond is not too long, animals perform on this task with an accuracy higher than chance.

Whereas the indirect realist approach would attribute correct responses to an internal representation of the past stimulus, in White's (1985) model of direct memory current responses are guided by the past stimulus itself. Put differently, in this model memory is nothing more than the direct perception of the past. White comments (p. 20) that his direct memory approach is informed by Gibson's (1979) rejection of the traditional distinction between perception and memory. Gibson's theory of direct perception can in turn be traced back to Holt's neorealism (Heft, 2001).[2] Here I propose a reevaluation of neorealism and an extension of the direct realist outlook to the whole of behavior analysis, first from short-term to long-term memory (Wilcox & Katz, 1981), and then to all conscious contents.

Reevaluating Neorealism

Holt's (1914) neorealism is subject to various objections (e.g., Rogers, 1920), and not all of them arise from misunderstandings or errors of logic. Any attempt at strengthening behavior analysis with neorealist concepts should thus avoid the mistakes of the past (Ryan, 2008). When examining Holt's neorealism it is particularly important to decide on what is dispensable and what is not. This cannot be done without some degree of philosophical risk-taking, but some of the options opened by neorealism seem more promising than others and therefore are worth exploring first.

The most fundamental concept of neorealism is that of cross-section. Basically, this concept designates a part of the world that is related in some way to another entity. The portion of the ocean illuminated by a navigator's searchlight, for example, is a cross-section (Holt, 1914, p. 171). In this case a part of the world (a part of the ocean) is related to another entity (the navigator) by the relation of being-illumined-by. Notice that the relation that defines the cross-section is selective, not creative: the searchlight merely selects, but does not create, the part of the ocean that it illuminates.

Holt (1914) assumed that the content of a person's consciousness was a cross-section, a part of the world related in some way to this person. This assumption fits our intuitive view of what perception consists of. When I look through the window of my office, for example, I am aware of a portion of the city of Braga, not the whole of it. Even when looking at what I take to be a single object, such as my desk, I am not aware of the totality of this object but only of a part of it (its top side for example). Finally, the part of which I am aware includes some, but not all, of the properties of this object. A cross-section is a part of the world but not a chunk, a point that Holt (1912) emphasized when he railed against the "crude brickbat notion of physical object" (p. 371). Hence, it is no objection to the neorealist identification of consciousness with its object to point out that

[2] In fact Gibson was one of Holt's students. Another well-known student of Holt was Edward C. Tolman. Although his molar behaviorism reflects the influence of Holt's (1915a, 1915b) relational view of cognition, the neorealist theory of conscious qualities is mentioned in Tolman's *Purposive Behavior in Animals and Men* (1932) only briefly and in a noncommittal fashion (p. 427).

fire burns whereas the idea of fire does not. As Holt (1912) retorted, fire burns but the shape of a flame does not burn either (p. 371).

Now if a person's consciousness is a cross-section defined by some relation to this person, we must decide on the nature of this relation. Causation, or at least causation with respect to a particular type of effect, is the most likely candidate (Montague, 1912). It is some causal relations between the properties of the top of my desk and me that make me see the top of my desk. In this case the causal medium is light. Blocking these causal relations by blocking light screens the object from view. But a causal relation from an object to the eyes, or even the brain, does not guarantee consciousness since a man in coma is unaware of his surroundings no matter how brightly illuminated. What prohibits the comatose patient from being conscious of the objects that surround him? Not a causal relation from the objects to his retina or even from the objects to his brain. What is missing in this case, I suggest, is a causal relation from the environment to the entirety of the patient's behavioral dispositions (Holt's "specific response").

This formulation of the causal criterion for consciousness is a working hypothesis (Tonneau, 2004). It may fail, but certainly it is worth taking as a starting point. The concept of behavioral disposition is used so as to avoid two horns of a dilemma. On the one hand, as the case of the comatose patient shows, we do not want to count as a criterion of consciousness any state of a neural receptor or even any neural state. We rather expect a person's consciousness to be available to behavior; conscious contents typically can be reported verbally and are able to influence a variety of actions (Holt, 1915b). On the other hand, we do not want to link consciousness to actual behavior too closely, since a person can be conscious without showing it at a behavioral level (dreaming being a good example of this). Requiring a person's consciousness to be related to a change of behavioral dispositions strikes the right balance.[3]

So far I have not proposed anything that would imply radical changes in Holt's (1912, 1914) original program. The most serious problem with neorealism, however, was its lack of plausibility with respect to hallucinatory experience. How could any theory which identifies consciousness with a cross-section of the environment explain cases of perceptual experience in which the object of perception is missing? Holt's (1912) solution was to embrace a metaphysics in which contradictions as well as unreal conscious contents were allowed to find their place: "The picture which I wish to leave is of a general universe of being in which all things physical, mental, and logical, propositions and terms, existent and non-existent, false and true, good and evil, real and unreal *subsist*" (p. 372). In Holt's metaphysical scheme, reality coincided not with the whole of the universe but only with a restricted portion of it, the portion "most remote from contradiction" (p. 366). The critics of neorealism pounded on Holt's treatment of error and existence (e.g., Rogers, 1920, p. 133).

Let us not dismiss Holt's (1912) Meinongian metaphysics too quickly, however. Holt's distinction between existence and subsistence can be restated in more restrained philosophical terms. Among the philosophers who are not nominalists about properties, some consider that properties are *universals*; that is, entities that can be instantiated by

[3] The price to be paid, of course, is that an adequate theory of dispositions must now be formulated. However difficult, this program is worth embarking on anyway, because dispositional concepts are ubiquitous in science regardless of the issues of consciousness and direct realism (Prior, 1985).

different objects. Armstrong (1997) has further proposed that there are no uninstantiated universals; universals must be instantiated by at least one object in order to be. Yet there may be good philosophical reasons to reject this doctrine. Bird (2007), for example, accuses Armstrong of a "failure of nerve" with respect to universals: "if universals really are entities in their own right, why should their existence depend upon a relationship with existing particulars?" (p. 55). The fact of hallucinatory experience may be a further reason to postulate universals that are not instantiated, since in hallucination we may be directly aware of them. Indeed, this seems to have been Holt's position. More recently, Johnston (2004) has argued that in hallucination we are acquainted with uninstantiated complexes of properties (such as the nonexistent shade of blue on my office desk).

The notion that perception without existent objects actually involves uninstantiated universals is attractive in its simplicity. I nevertheless suggest that it should be resisted as much as possible. A first problem with uninstantiated properties is their apparent lack of location in space. An instantiated universal may be present in all of its instances (Donagan, 1963), but uninstantiated universals do not have this privilege. The problem may not be fatal, but it should not be dismissed lightly either, since spatial location may be the mark of the physical (Markosian, 2000). Admitting spatially nonlocated entities along with others that are located in space comes dangerously close to admitting a bifurcated ontology of the physical and the nonphysical. A second, related, problem is that uninstantiated universals seem disconnected from causation (Armstrong, 1997), which we have previously assumed to be part of the defining relation for the cross-section of consciousness. It is hard to see how nonexistent properties could be causally related to someone's actions or how a person could "respond" to them (to use Holt's term). Finally, even if we admit absences in our ontology (e.g., Kukso, 2006), in the nonveridical perception of a property profile we do not seem to be aware of the *absence* of this profile; on the contrary, we seem to be aware of its very presence in consciousness.

For all three reasons, I suggest that a direct realist account of nonveridical perception should be found elsewhere than in Holt's appeal to uninstantiated properties. An alternative is to identify hallucinations and dreams with profiles of properties instantiated at different moments in a person's history of interaction with the environment (Boag, 2008; Tonneau, 2004). Appealing to instantiated properties guarantees that conscious contents are located in spacetime and can play a causal role with respect to behavior. Finally, appealing to instantiations that are distributed in time fits with the remainder of behavior analysis (e.g., Hineline, 1990; Skinner, 1931) better than Skinner's (1974) own forays into indirect realism and unsuccessful appeals to internal acts of sensing.

Causation and Time

Behavior analysts are well aware of the fact that current performance depends on previous conditions, often remote in time from the action being measured. This sort of dependence is typically referred to as historical causation (Morris, Higgins, & Bickel, 1982) or action at a temporal distance (Marr, 1984). Some behavior analysts (e.g., Schaal, 2005) also refer to Russell's (1921) concept of mnemic causation, which includes "past occurrences in the history of the organism as part of the causes of the present response" (p. 78). But all of these concepts, although pointing in the right direction, are prone to

misunderstanding and need to be clarified.

First, as Marr (2008) correctly pointed out, all cases of historical causation with respect to behavior involve physiological mediation. When some past event E affects current behavior (B), we can safely assume that there is at least one continuous causal chain or stream leading from E to B, a causal stream located partly in the nervous system (Schaal, 2005). Hence, the sort of causation that behavior analysts need with respect to memory cannot be *unmediated* causation. Yet, to say that historical effects are causally mediated by the organism's physiology is not to say that historical causation is spurious. Russell's (1921) concept of mnemic causation apparently was meant to be provisional and was therefore ambiguous in this respect: "I do not wish to urge that this form of causation [mnemic causation] is ultimate, but that, in the present state of our knowledge, it affords a simplification, and enables us to state laws of behaviour in less hypothetical terms than we should otherwise have to employ" (p. 85). If the fate of the concept of mnemic causation hangs on considerations of simplicity and on our temporary ignorance of mediating mechanisms, as Russell suggested, then this concept might well disappear once the mediating mechanisms are worked out.

Clearly this is not the concept of causation that is needed in connection with our perception of the past. From the standpoint of direct realism, a person's awareness of past events must be grounded in a metaphysically robust relation to this person (as explained below), a relation impervious to the growth of scientific knowledge about the brain. This relation must be present in memory phenomena, and hold among the actual events that compose them, regardless of what future studies of physiology may or may not bring up.

The correct picture of this relation, as I see it, is this. An environmental event (A) causes a change in the nervous system (B). The modified neural state persists for a period of time and then causes a change of behavior (C), either spontaneously or under the influence of additional side conditions (e.g., the "retrieval cues" that are of interest to memory psychologists). Once completed, this process constitutes a causal chain from A to B to C, the persistence of B in the second stage of the chain being itself a causal chain.[4] Clearly we are not dealing with "action at a temporal distance" in the sense of causation from A to C *in the absence of any intermediate link*. For an intermediate link (B) *is* present. Its presence in the chain, however, does not imply that *A does not cause C at all*. The latter conclusion would hold if A and B were competitors and if the causal relation from B to C prevented any causal relation from A to C. But we have no reason to assume that A and B are competitors in this sense (Shapiro & Sober, 2007, p. 260). Far from *preventing* the causal relation from A to C, the existence of an intermediate link in the causal chain (B) *causes* A to cause C. In a nutshell, A causes C, because A causes B and B causes C. Causation being arguably transitive,[5] as causal chains extend across time, they allow the sphere of influence of their initial link to reach ever more distant events.

---

[4] Armstrong (1997) calls this sort of process, in which the state exemplified at a given time causes its exemplification at later times, "immanent causality" (p. 73).

[5] Philosophers have proposed various counterexamples to the transitivity of causation. In all cases that I am aware of, however, either transitivity does hold (even though it may be awkward to acknowledge, as when one's birth is said to be a cause of one's death) or the example has been misdescribed as involving a common intermediate link (B) when in fact it does not. See Paul (2000) for further discussion.

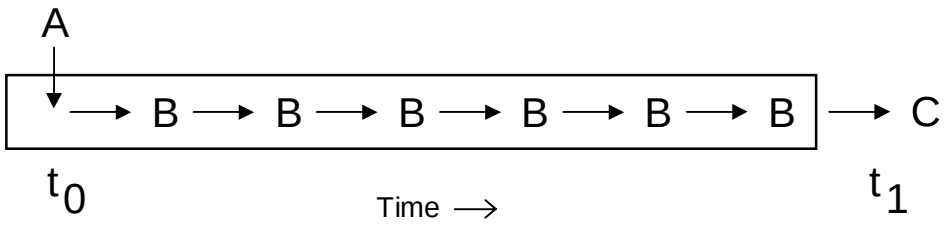

Figure 1. The response mechanism in memory. The person, represented by a rectangle, reacts to an event A with a continuous causal chain that includes events in the brain (B) and eventuates in a change of behavior (C). Time flows from left to right.

We can therefore acknowledge that the causal relation in memory involves two events A and C separated in time (hence the grain of truth in "action at a temporal distance") while depending on the existence of mediating links (hence the truth in the notion of physiological mediation) but remaining as real as any causal relation between two contiguous events (hence the truth of direct realism about the past). It is this actual causal relation from A to C which grounds the phenomenon of memory, as shown in Figure 1. In this figure time flows from left to right. The rectangle represents a person (or animal) extended in time and in direct contact with the event A at time t_0. The causal chain that starts with A, continues with B, and eventuates in C at time t_1, is nothing less than the person's mechanism of response to A, this response mechanism being itself extended in time. Notice that the person has no need to "look back" or "travel backward in time" in order to perceive A. She can look at A directly because her *eyes* are extended in time!

A theory of memory in terms of direct perception of the past requires a conception of causation that is singularist (e.g., Anscombe, 1981) and non-Humean (Montague, 1912). Aside from their failure to provide a satisfactory analysis of causal notions (a failure of which philosophers are becoming increasingly aware), Humean approaches to causation in terms of statistical regularities are especially inappropriate in a case such as that of Figure 1. Of the many events and features present in the environment at time t_0, all of them equally correlated with B and C, only one of them (A) causes B and then C. Causation takes place on a single-case basis and therefore cannot be reduced to co-location, contiguity, correlation, or pairing.

The view of memory illustrated in Figure 1 is obviously inconsistent with presentism (Dainton, 2001), defined as the philosophical belief that only the present exists (in some suitably restricted sense of "present"). But there are good reasons to reject presentism anyway; retaining the possibility of explaining our experience of the past as experience of the past itself is one of them. Acknowledging the reality of relations across time also defuses an alleged inconsistency in the notion of the direct memory. Drake (1917), for example, criticizes the concept of direct memory on the following basis: "the impossibility that a sense-quality existing *now* in my experience should be numerically identical with any aspect of an object which no longer exists is so obvious that the argument needs no emphasis" (p. 371). In fact, such a numerical identity is not only

possible, it is mundane if direct realism is correct. In the case of Figure 1, my experience at t_1, being identical with A, is indeed located at t_0. What makes it my experience at t_1 is not its temporal location (t_0) but its causal *relation* to my behavior at t_1.

Analogously, consider the spatial relation between my neighbor, Jim, and myself. What makes him *my* neighbor is his relation to the house where I live. In order to satisfy this relation, however, my neighbor does not need to be me nor live in my house. He and I just exist at different places (call them s_0 and s_1), and the spatial relation between s_0 and s_1 causes him to be my neighbor. Even if we divide up the space between us into several spatial intermediaries, and though there is a fence between us, no one would argue that it is the fence that is my neighbor, not Jim. Claiming, as Drake (1917) does, that my experience at t_1 cannot be located at t_0 is like claiming that my neighbor cannot live in the house next to mine. Of course he can, and he does. In the spatial case no less than in the temporal case, my experience at t_1, or my neighbor here in Braga, qualifies as such by virtue of a relation that it entertains to *another* entity or event, located at another moment or another place. There is no absurdity, and Drake's claim—no matter how intuitive it may seem—should be rejected.

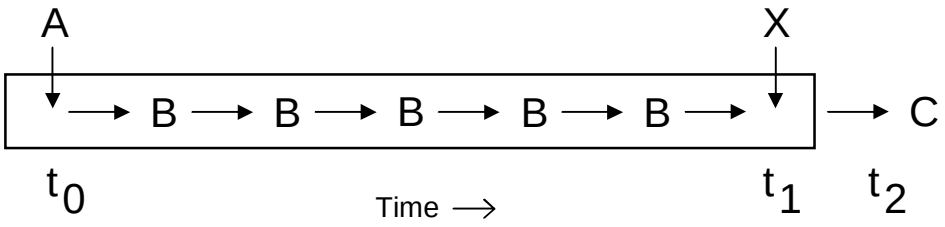

Figure 2. Context-dependent memory. The event present at t_0 affects behavior at t_2 provided that a contextual cue X is present at t_1. Same conventions as in the previous figure.

One last issue is worth discussing. Consider Figure 2, in which an event A present at t_0 influences behavior at t_2, but only if a contextual cue X is present at t_1 (hence, *after* t_0). I have argued elsewhere (Tonneau, 1990) that conceptualizing direct memory in terms of momentary events lead one to rely on backward causation in the case of Figure 2. The argument I offered was that X needed to act backward in time from X to A (hence from t_1 to t_0) in order to trigger a causal relation between A and C (hence between t_0 and t_2). To eliminate this difficulty, I recommended switching to a conception of direct memory in terms of temporally extended patterns instead of isolated events; in Figure 2, for example, I proposed that the environmental cause of C was neither A nor X but the molar sequence, A-X located at (t_0, t_1).

Evaluating the possible roles of molar variables such as sequences, densities, and rates of events is an important task for behavior analysis (Tonneau, 2005b), and I do not deny that we can react to molar sequences in some circumstances. But I now think that veridical memory is not one of these and that my 1990 argument was misguided. In the case of Figure 2, and contrary to what I thought, X has no need to act on A in order to cause

A to cause C. All that X needs to do instead is to act on *B* and thereby complete a causal chain that was already under way. Assuming again that causation is transitive, the fact that A causes B and B causes C now ensures that A causes C (Figure 2).

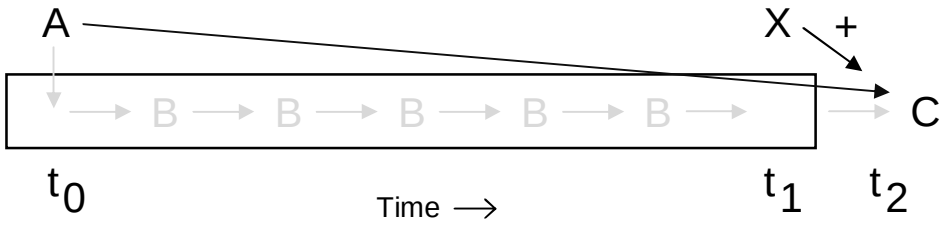

Figure 3. Context-dependent memory and the causal relations between environment and behavior. The solid arrows represent the causal relations that are made possible by mediating physiological links (shown in gray). Same conventions as in the previous figure.

Focusing on the causal relations between environment and behavior instead of the physiological makeup that makes these relations possible, the picture is that of Figure 3. The contextual cue present at time t_1 (X) causes a causal relation between A (at time t_0) and C (at time t_2).[6] Following Donagan (1963) and assuming, as we do, that relations are located where their bearers are located, the causal relation between A and C that is caused by X occupies the pair (t_0, t_2), whereas X occupies t_1. The higher-order causal relation, indicated by a plus sign, between X and the causal relation between A and C goes from t_1 to (t_0, t_2). There is no causal arrow from t_1 to t_0, which would be the case if Figure 3 involved backward causation in a traditional sense. Now if the existence of a causal arrow from t_1 to (t_0, t_2) is deemed to involve backward causation because the relation between A and C starts before t_1, then so be it; but it is backward causation of a benign sort, as when repeating the end of an improvised melody makes its first note more distant from the last. Only the relational properties of the first note are affected.

The Prevalence of Time

A neorealist perspective on memory would be of little importance to behavior analysis if the issues raised by the direct perception of the past were limited to a few isolated phenomena. In fact historical causation of the sort illustrated in Figure 3 is pervasive and extends to all corners of behavior analysis, some of them well explored

[6] Again, keep in mind that the causal relations from A to B, B to B, and B to C (shown in gray) do not compete with the causal relation from A to C (shown in black). On the contrary, the former relations are necessary for the latter to take place. The arrow notation for causal relations easily leads to misunderstandings in this respect. So does my use of a discrete notation (A, X, etc.) for stimuli, which conveys the unwarranted picture of a cross-section made of objects coinciding with simple physical boundaries. Finally, Figure 3 neglects the feedback loops that are involved in the maintenance and dynamics of behavior (e.g., Baum, 1973).

already, many of them ripe for study. Consider Pavlovian conditioning, in which the behavioral effects of the conditional stimulus (CS) change due to its previous correlation with an unconditional stimulus (US). There is nothing illogical in assuming that a regimen of CS-US correlation can cause the CS to cause a conditional response (CR), and indeed this is how Pavlovian conditioning is commonly discussed in behavior analysis. This description assumes a hierarchical arrangement of causation with a first-order causal relation (from the current CS to the CR) caused by a second-order variable (the previous CS-US correlation).

A problem with this hypothesized causal structure is that it fails to account for the well-established phenomena of US devaluation and revaluation (e.g., Rescorla & Holland, 1982), in which post-conditioning manipulations of the unconditional stimulus affect the responses evoked by the US and the CS in a parallel fashion. Revaluation phenomena are unpredicted from a standard behavior-analytic standpoint because the observed changes in the CR take place in the absence of further CS-US pairings, hence in the absence of "conditioning" as the term is usually defined.

Behavior analysts can of course incorporate US revaluation in their account of Pavlovian conditioning by adding to it an *ad hoc* principle to the effect that such phenomena do take place. Nevertheless, US revaluation effects may be explained in terms of more basic principles. Some cognitive psychologists have proposed for example that Pavlovian conditioning involves a network of associations and that one of these consists of a link between representations of the CS and the US (Holland, 1990). When the CS is presented, activation runs from the CS node to the US node, which then generates the conditional response. This network model explains why the conditional and unconditional responses are similar: both involve the activation of the US node. Revaluation effects are explained by the same mechanism. Through revaluation, the behavioral significance of the US node is modified, but presentation of the CS activates the US node as before, resulting in a modified conditional response. In sum, in this type of associative memory model the CR reflects the behavioral properties of the US because the CR is evoked by a *representation* of the US (Holland, 2008).

The same facts can be accounted for by assuming that in Pavlovian conditioning the conditional response is caused, not by a representation of the US, but by the US itself. From a direct realist standpoint, it is the US itself (playing the role of A in Figure 3) which causes the conditional response when the CS (playing the role of a contextual cue) is presented. This sort of memory phenomenon can be described by Figure 3 with A as the unconditional stimulus, X as the conditional stimulus, and C as the conditional response. The only ingredient of Pavlovian conditioning that is missing from Figure 3 is the history of pairings between X and A which allows X to cause a causal relation between A and behavior. Including this history in the explanatory account leads us to a three-tier causal hierarchy in which (1) a history of correlation between X and A causes (2) the causal relation between X and (3) the causal relation between A and C. The direct realist account is analogous to a memory model with an actual correlation in lieu of the associative link, and actual events (remote in time from the current response) in lieu of internal representations.

The context dependence illustrated in Figure 3 is ubiquitous with respect to behavior, from the conditional responses observed in Pavlovian settings to complex cognitive performance (such as reports of memory or imagery in humans). The neorealist

explanation of these phenomena requires identifying the part of the environment that is causally related to current behavior (Holt, 1914). Thus, neorealism may look like an old stimulus-response psychology. And to some extent it is; only that the "stimulus" responsible for behavior is typically remote from current responding. Assume for instance that at time t_1 John has a mental image of the house in which he lived as a child (at time t_0). Aside from neorealism, virtually *all* psychological accounts of this phenomenon (including Skinner's) identify John's mental image with a perceptual activity located somewhere in John's body at time t_1. By contrast, neorealism identifies John's mental image of his past house with his past house itself, or at least a section of it, at time t_0. It is not John who is seeing his house; rather, it is his house that shows itself to him (Tonneau, 2004).

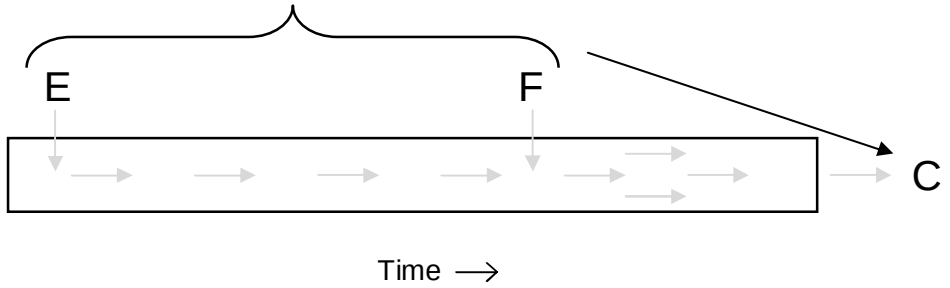

Figure 4. Perceiving disconnected parts of the environment. In this example a pattern made of two disconnected parts, {E, F}, acts on one person (represented as a rectangle and extended in time). Same conventions as in the previous figure.

Of course, our visual memories of a past object may never be exactly identical to one another or indeed to this object itself. But this is no objection to neorealism, since, as we have seen, the cross-section active at any moment will include only a subset of features of the relevant object or of the environment of which it is part. *Which* subset will depend on the state of the person at this moment.[7] Because this state is in permanent flux, the person's cross-section will shift its content from one portion of the environment to another (Tonneau, 2004). In more complicated cases, the person presumably reacts to disconnected parts of the environment, as in Figure 4, in which a disconnected pattern such as {E, F} causes the current response. As in previous figures, the neorealist account sketched in Figure 4 requires a singularist concept of causation, since all of the individual components in the environment, such as {E} and {F}, as well as {E, F}, are equally correlated with responding.

The hypothesis that people can react to components that are scattered in time, points to what I think is an adequate realist solution to the puzzle of hallucinations and dreams

[7] Which is not to say that *the* subset in question depends on the person, still less that it resides inside her. Concluding otherwise would be like arguing that the walls I see must be in my neck because the position of my neck determines *which* walls I see. For a more detailed treatment of this issue, see Montague (1912) and Tonneau (2004). Other antirealist fallacies (some of them, unfortunately, endorsed by behavior analysts themselves) are addressed by Tonneau (2005a).

(Tonneau, 2004). A complete analysis is obviously beyond the bounds of this essay, but the nature of the correct solution is not beyond all conjecture and can be sketched easily. The person who is aware of a conscious content with the properties of a blue horse cannot be aware of a blue horse, since there are no blue horses in the environment. Like {E, F} in Figure 4, however, the disconnected pair {blue, horse} *is* present in the temporally extended environment of anyone who has met at least one blue thing and at least one horse. Except for the property of being disconnected, all of the perceptible features of this pair are features that a blue horse would have. Accordingly, in the version of neorealism that I am proposing, this pair *is* the blue-horse-like content. That is, I am proposing that a person who claims to "see a blue horse" is, in fact, responding to some blue thing *and* some horse in her past (not responding to some blue thing *and* responding to some horse, mind you, but responding to the *pair*, {blue, horse}). There is nothing "unreal" to which she is responding, just patterns that are extended in time and to which a third party does not have easy access.

A common objection is to remark that the blue-horse-like content that can appear in one's hallucinations does not look like a blue spot and a horse side by side. The objection, however, misrepresents the situation that holds in nonveridical perception. When I look at a blue spot and a horse side by side, I react not only to the spot and the horse but also to the stretch between them, a stretch that is neither blue nor horse-like (Montague, 1912). A better analogy to hallucination would be to react to the blue color of the spot and the shape of the horse without reacting to the stretch between them. Transposed along the time dimension, the person subjected to a hallucination reacts to disconnected components of the environment while failing to react to their disconnectedness. Hallucination may be a case of faulty perception, but contra popular theories its only "fault" lies in not including enough of the environment instead of including more than what there is.

Conclusion

Direct and indirect realism (and even forms of antirealism) coexist uncomfortably in current operant theory (Tonneau, 2005a). In a single essay, for example, Branch (1987) can defend both an interpretation of mental events in terms of historical causation and an interpretation in terms of covert sensory activities. I support the former and reject the latter on philosophical and ontological grounds. The distinguishing feature of behavior analysis lies in its concentration on relations between environment and behavior, a subject domain shared with ecological psychology. An interpretation of perception in terms of covert sensing activities cannot be integrated with the remainder of the discipline. Furthermore, the indirect realist approach to perception propagates from nonveridical to veridical conscious contents so as to jeopardize the very evidence that is adduced in its support (Katz & Frost, 1979).

If my argument is correct, the only way to escape from this impasse is a return to neorealism as a philosophy of perception and consciousness (Tonneau, 2004). Whereas White's (1985) direct realist approach focuses on delays of a few seconds, or at best a few minutes due to the nature of the match-to-sample task, I propose to extend the direct perception approach to long-term remembering and from then to imagery, hallucinations, and dreams in human subjects. Needless to say, the validity of the neorealist program

depends on the possibility of bringing it to fruition through a combination of conceptual reflection, theory, and empirical research.

Behavior analysts are unlikely to be swayed by philosophical speculation, but to repeat, those who neglect the philosophical issue of direct realism do so to their own empirical peril. Psychophysics, long-term remembering, and perceptual plasticity (Goldstone, 1998), are only a sample of connected issues that call neither for dissolution nor neglect but rather for integration with Skinner's operant concepts. At the very least, the study of such phenomena could provide behavior analysts with a better understanding of how selected parts of the environment become available for behavioral control. At best, the study of memory from a direct realist standpoint could merge into a comprehensive research program on veridical as well as nonveridical perceptual content and consciousness (Katz & Frost, 1979). Behavior analysts and ecological psychologists may be more aware than before of the commonalities between their respective approaches (Costall, 1984; Morris, 2009). It is time for behavior analysis to come back to its realist roots (O'Neil, 1968) and face the hard issue of perception.

References

Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences, 14*, 471-517. (Includes commentary.)

Anscombe, G. E. M. (1981). Causality and determination. In G. E. M. Anscombe (Ed.), *The collected philosophical papers of G. E. M. Anscombe: Vol. 2. Metaphysics and the philosophy of mind* (pp. 133-147). Minneapolis, MN: University of Minnesota Press.

Armstrong, D. M. (1997). *A world of states of affairs*. New York: Cambridge University Press.

Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior, 20*, 137-153.

Bird, A. (2007). *Nature's metaphysics: Laws and properties*. New York: Oxford University Press.

Blough, D., & Blough, P. (1977). Animal psychophysics. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 514-539). Englewood Cliffs, NJ: Prentice-Hall.

Boag, S. (2008). 'Mind as feeling' or affective relations? A contribution to the school of Andersonian realism. *Theory & Psychology*, *18*, 505–525

Branch, M. N. (1987). Behavior analysis: A conceptual and empirical base for behavior therapy. *Behavior Therapist, 4*, 79-84.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Chomsky, N. (1959). [Review of the book *Verbal Behavior*]. *Language, 35*, 26-58.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*, 7-19.

Cohen, D. J., & Blair, C. (1998). Mental rotation and temporal contingencies. *Journal of the Experimental Analysis of Behavior,* 70, 203-214.

Costall, A. P. (1984). Are theories of perception necessary? A review of Gibson's *The Ecological Approach to Visual Perception*. *Journal of the Experimental Analysis of Behavior, 41*, 109-115.

Dainton, B. (2001) *Time and space*. Montreal: McGill-Queen's University Press.

Davison, M. C., & Tustin, R. D. (1978). The relation between the generalized matching law and signal-detection theory. *Journal of the Experimental Analysis of Behavior, 29*, 331-336.

Donagan, A. (1963). Universals and metaphysical realism. *Monist, 47*, 211-246.

Drake, D. (1917). A cul-de-sac for realism. *Journal of Philosophy, Psychology and Scientific Methods*, *14*, 365-373.

Gibson, J. J. (1960). The concept of the stimulus in psychology. *American Psychologist, 15*, 694-703.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585-612.

Hayes, L. J. (1994). Thinking. In S. C. Hayes, L. J. Hayes, M. Sato & K. Ono (Eds.), *Behavior analysis of language and cognition* (pp. 149-164). Reno, NV: Context Press.

Heft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism*. Mahwah, NJ: Erlbaum.

Hineline, P. N. (1990). The origins of environment-based psychological theory. *Journal of the Experimental Analysis of Behavior, 53*, 305-320.

Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition, 37*, 105-131.

Holland, P. C. (2008). Cognitive versus stimulus–response theories of learning. *Learning & Behavior, 36*, 227-241.

Holt, E. B. (1912). The place of illusory experience in a realistic world. In E. B. Holt, W. T. Marvin, W. P. Montague, R. B. Perry, W. B. Pitkin & E. G. Spaulding (Eds.), *The new realism: Coöperative studies in philosophy* (pp. 303-373). New York: Macmillan.

Holt, E. B. (1914). *The concept of consciousness*. New York: Macmillan.

Holt, E. B. (1915a). Response and cognition. I. The specific-response relation. *Journal of Philosophy, Psychology and Scientific Methods*, *12*, 365-373.

Holt, E. B. (1915b). *The Freudian wish and its place in ethics*. New York: Henry Holt.

Holt, E. B., Marvin, W. T., Montague, W. P., Perry, R. B., Pitkin, W. B., & Spaulding, E. G. (1910). The program and first platform of six realists. *Journal of Philosophy, Psychology and Scientific Methods*, *7*, 393-401.

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36*, 405-445.

Johnston, M. (2004). The obscure object of hallucination. *Philosophical Studies*, *120*, 113–183.

Katz, S., & Frost, G. (1979). The origins of knowledge in two theories of brain: The cognitive paradox revealed. *Behaviorism, 7*(2), 35-44.

Kukso, B. (2006). The reality of absences. *Australasian Journal of Philosophy*, *84*, 21-37.

Malone, J. C., Jr. (1978). Beyond the *operant* analysis of behavior. *Behavior Therapy*, *9*, 584-591.

Markosian, N. (2000). What are physical objects? *Philosophy and Phenomenological*

Research, *61*, 375-395.
Marr, M. J. (1984). Conceptual approaches and issues. *Journal of the Experimental Analysis of Behavior*, *42*, 353-362.
Marr, M. J. (2008). The abdication of belief: A comment on Foxall's replies to his critics. *Behavior and Philosophy*, *36*, 157-168.
Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology*, *19*, 329–344.
Montague, W. P. (1912). A realistic theory of truth and error. In E. B. Holt, W. T. Marvin, W. P. Montague, R. B. Perry, W. B. Pitkin & E. G. Spaulding (Eds.), *The new realism: Coöperative studies in philosophy* (pp. 251-300). New York: Macmillan.
Morris, E. K. (2009). Behavior analysis and ecological psychology: Past, present, and future. A review of Harry Heft's *Ecological Psychology in Context*. *Journal of the Experimental Analysis of Behavior*, *92*, 275–304.
Morris, E. K., Higgins, S. T., & Bickel, W. K. (1982). Comments on cognitive science in the experimental analysis of behavior. *Behavior Analyst*, *5*, 109-125.
Nevin, J. A. (1981). Psychophysics and reinforcement schedules: An integration. In M. L. Commons & J. A. Nevin (Eds.), *Quantitative analyses of behavior: Vol. 1. Discriminative properties of reinforcement schedules* (pp. 3-27). Cambridge, MA: Ballinger.
Nevin, J. A., Davison, M., & Shahan, T. A. (2005). A theory of attending and reinforcement in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 84, 281-303.
Newell, A. (1980). Physical symbol systems. *Cognitive Science*, *4*, 135-183.
O'Neil, W. M. (1968). Realism and behaviorism. *Journal of the History of the Behavioral Sciences*, *4*, 152-160.
O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*, 939–1031. (Includes commentary.)
Paul, L. A. (2000). Aspect causation. *Journal of Philosophy*, *97*, 235–256.
Prior, E. (1985). *Dispositions*. Aberdeen, Scotland: Aberdeen University Press.
Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.
Pylyshyn, Z. W. (1987). What's in a mind? *Synthese*, *70*, 97-122.
Rachlin, H. (1973). Contrast and matching. *Psychological Review*, *80*, 217-234.
Rachlin, H. (1991). The cognitive laboratory, the library and the Skinner box. *Behavioral and Brain Sciences*, *14*, 501.
Rachlin, H. (1994). *Behavior and mind: The roots of modern psychology*. New York: Oxford University Press.
Rescorla, R. A., & Holland, P. C. (1982). Behavioral studies of associative learning in animals. *Annual Review of Psychology*, *33*, 265-308.
Roediger, R. (2004). What happened to behaviorism? *APS Observer*, *17*(5), 40–42.
Rogers, A. K. (1920). The problem of error. In D. Drake, A. O. Lovejoy, J. B. Pratt, A. K. Rogers, G. Santayana, R. W. Sellars & C. A. Strong (Eds.), *Essays in critical realism: A co-operative study of the problem of knowledge* (pp. 117-160). London: Macmillan.

Russell, B. (1921). *The analysis of mind*. London: George Allen & Unwin.

Ryan, F. X. (2008). Neo-realism. In J. Lachs & R. Talisse (Eds.), *American philosophy: An encyclopedia* (pp. 541-542). New York: Routledge.

Schaal, D. W. (2005). Naming our concerns about neuroscience: A review of Bennett and Hacker's *Philosophical Foundations of Neuroscience*. *Journal of the Experimental Analysis of Behavior, 84*, 683–692.

Shapiro, L., & Sober, E. (2007). Epiphenomenalism: The do's and the don'ts. In P. Machamer & G. Wolters (Eds.), *Thinking about causes: From Greek philosophy to modern physics* (pp. 235-264). Pittsburgh, PA: University of Pittsburgh Press.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127-146.

Skinner, B. F. (1931). The concept of the reflex in the description of behavior. *Journal of General Psychology, 5*, 427-458.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1972). What is psychotic behavior? In B. F. Skinner (Ed.), *Cumulative record: A selection of papers* (pp. 257-275). New York: Appleton-Century-Crofts.

Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.

Skinner, B. F. (1989. The origins of cognitive thought. *American Psychologist, 44*, 13-18.

Smith, T. L. (1988). Neo-Skinnerian psychology: A non-radical behaviorism. In A. Fine & J. Leplin (Eds.), *PSA 1988* (Vol. 1, pp. 143-148). East Lansing, MI: Philosophy of Science Association.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Century.

Tonneau, F. (1990). From reflex to memory: Molar sequences in Pavlovian and instrumental conditioning. *Psychological Record, 40*, 587-607.

Tonneau, F. (2004). Consciousness outside the head. *Behavior and Philosophy, 32*, 97-123.

Tonneau, F. (2005a). Antirealist arguments in behavior analysis. *Behavior and Philosophy, 33*, 55-65.

Tonneau, F. (2005b). Windows. *Behavioural Processes, 69*, 237-247.

Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn. *Cognition, 9*, 237-304.

White, K. G. (1985). Characteristics of forgetting functions in delayed matching to sample. *Journal of the Experimental Analysis of Behavior, 44*, 15-34.

White, K. G. (1991). Psychophysics of direct remembering. In J. A. Commons, M. C. Davison, & J. A. Nevin (Eds.), *Models of behavior: Signal detection* (pp. 221–237). New York: Erlbaum.

White, K. G., & J. T. Wixted (1999). Psychophysics of remembering. *Journal of the Experimental Analysis of Behavior, 71*, 91-113.

Wilcox, S., & Katz, S. (1981). A direct realistic alternative to the traditional conception of memory. *Behaviorism, 9*, 227-239.

Williams, J. L. (1974). Evidence of complementary afterimages in the pigeon. *Journal of the Experimental Analysis of Behavior, 21*, 421-424.

Zuriff, G. E. (1972). A behavioral interpretation of psychophysical scaling. *Behaviorism, 1*, 118–133.