INTENTIONAL CONSEQUENCES OF SELF-INSTRUCTION

Gordon R. Foxall
Cardiff University

Jorge M. Oliveira-Castro
Universidade de Brasilia

ABSTRACT: Discrepancies between animal and human responding on standard schedules of reinforcement have been explained by reference to the human capacity for language and consequent formulation of self-instructions. As a result, schedule responding has been causally attributed to private events. However, the operations that individuals are assumed to carry out in the formulation of self-instructions cannot be described other than intentionally and this raises important issues of explanation for an extensional behavioral science. It is argued that radical behaviorism is ultimately dependent upon intentional explanation; moreover, the ascription of causality to the intentional terms on which radical behaviorism is dependent leads to the necessity of incorporating cognitive explanation into operant psychology. Extensions of radical behaviorist explanation beyond accounts of contingency-shaped behavior cannot avoid the use of intentional terms.

Key words: intentionality, scheduled contingencies, self-instruction, verbal behavior, behavioral explanation

Radical behaviorism has progressed not only through the accumulation of data but by conceptual reformulations that transformed both its ontological purview and its methodological practice. Its origin as a separate theoretical position within psychology stems from the adoption of a novel theoretical stance that has had pervasive implications for both these concerns: the construal of behavior and its causes and the scientific means of producing data consistent with this view. Skinner’s (1938, 1953) delineation of the field he proposed rested importantly on the ontological distinctions he drew between respondent conditioning based on “involuntary” reflexive learning of elicited responses, and operant conditioning in which “voluntary,” emitted behavior comes under the control of its consequences. His unique approach to the analysis of behavior relied on his ability to “make over the entire field,” to persuade other psychologists of

AUTHORS’ NOTE: Please address correspondence to Gordon Foxall, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, Wales, United Kingdom; Email: foxall@cf.ac.uk.

While Skinner generally avoided terms such as “voluntary” on account of their implication of personal control of behavior, he expressly employed this term in Science and Human Behavior (1953, pp. 110-113), noting, however, that “voluntary” behavior, as he defined it, is no more free of contingency control than is “involuntary.”
this crucial distinction and the necessity of an innovative methodology by which to define the key variables and to empirically investigate and explain operant response (Skinner, 1931, 1935a, 1935b, 1937, 1950, 1956, 1957)\textsuperscript{2}. The contribution of both fact-gathering and conceptual redefinition to the growth of knowledge is not, of course, peculiar to radical behaviorism: it is a central component of all active research programs that over time come to delineate distinctive explanatory perspectives (Kuhn, 1970; Lakatos & Musgrave, 1970; Machado & Silva, 2007).

The essential characteristic of radical behaviorism is the explanation of behavior in terms of Machian positivism (i.e., the description of functional relationships between environment and responding that lend themselves to extensional language, avoiding reference to intentionality; Skinner, 1945; cf. Dennett, 1969; Foxall, 2004; L. D. Smith, 1986; T. L. Smith, 1994; Taylor, 1964). Cognitive psychology, by contrast, is founded upon intentional reasoning, attributing behavior to beliefs and desires, attitudes and intentions, among other mental events and processes. This means not only that it seeks to predict and explain observed behavior in terms of these intentional terms, but that it takes on board the implications of intentionality in its philosophical sense which refers to the \textit{aboutness} of these idioms. Beliefs, desires, attitudes, and intentions are inevitably \textit{about} something other than themselves. It is this referential characteristic of these psychological attitudes that constitutes their intentionality in the philosophical sense, and it is this meaning of \textit{intentionality} which is central to this paper (see, inter alia, Anscombe, 1957; Chisholm, 1957; Dennett, 1996; Quine, 1960; Russell, 1912; Searle, 1983). To use intentional idioms is to partake in a form of explanation that is distinct from the extensional approach of radical behaviorism, but to do so is not, of itself, to embrace “cognitivism” or even “creationism”\textsuperscript{3}. However, were radical behaviorist explanation at some point to yield to the necessity of incorporating intentionality, some serious rethinking of the philosophical foundations of extensional behavioral science would be required.

Not only does radical behaviorism not share the intentionalistic approach of cognitive psychology: it is apparently unique among the neobehaviorisms in avoiding locutions of this kind (Foxall, 2004). Tolmanian behaviorism overtly resorts to the language of expectancy, and even Hullian psychology, usually presented as wholly extensional, cannot entirely do without reference to anticipation (Deutsch, 1960; Hull, 1952; Kitchener, 1977, 1979, 1996; MacCorquodale & Meehl, 1954). Expecting and anticipating fall into the intentional camp in that they are inherently about something other than themselves. Herein lies the philosophical definition of intentionality which has often not been part of psychological behaviorists’ use of the term. The elements of the n-term contingency are not of this kind. A discriminative stimulus is simply one in the presence of which learned behavior is more likely to be repeated; a response is an

\textsuperscript{2} The reference to making over the entire field comes from Skinner’s autobiography.

\textsuperscript{3} The scare quotes are unfortunately necessary since they denote the labeling that so often is taken to embody intellectual refutation.
emitted piece of behavior that may be repeated if it is followed by certain consequences; a reinforcer is a consequence of behavior that has the effect of making the response that produced it more probable in the future; a motivating event increases the salience of a reinforcer. The language of intentionality is not required to describe, predict, and control behavior in these terms. Interesting theoretical issues are raised, however, by the inclusion of verbal behavior in one’s repertoire (T. L. Smith, 1994; Foxall, 1999).

In 1969, Skinner made a further major theoretical advance, based on a distinction between contingency-shaped and rule-governed behavior, which has become fundamental to radical behaviorism as a philosophy of psychology (Skinner, 1969). It has given rise, especially in the last two decades, to a large volume of empirical and theoretical investigation. In particular, it has led to the finding that the behavior of verbal humans may show insensitivity to changing schedule requirements. The source of such insensitivity may be the instructions provided by another person (e.g., an experimenter). But it may also arise from rules devised by the experimental subject himself, regardless of both the actual contingencies and the instructions provided by the experimenter. It is at this point that our theoretical interest is piqued. We argue in this paper that, as a result of the attempt to account for verbal behavior in operant terms and, specifically, to use verbal behavior to explain nonverbal responding, the uniqueness and independence of radical behaviorism as the philosophy of a wholly extensional behavioral science cannot be sustained. We first set out the facts regarding the non-contingent nature of rule-governed behavior, and then examine the explanations which some radical behaviorists have provided of this phenomenon in terms of self-derived instructions. This leads on naturally to the argument that their accounts require the adoption of intentionalistic language and that the calculative requirements of self-rule formulation that they assume amount to cognitive reasoning.

Instructed Schedule Performance

The phenomenon of interest has been recognized for several decades. Baron, Kaufman, and Stauber (1969) summarize work from the 1950s and 1960s on the effects of instructions on scheduled behavior. For instance:

. . .when instructions about the desired response are omitted, substantial numbers of subjects may fail to acquire the response despite scheduling of reinforcing contingencies deemed favorable for acquisition. (p. 701)

Indeed, there is strong evidence that unless intentional concepts and the variables derived from them specify situational context, i.e. rest upon the contextual stance (Foxall, 1999), they prove unreliable means of prediction and thus potentially of control (Foxall, 2004, 2005).

In fact, Skinner’s original paper was published in 1966 but, as is the common practice, reference is given here to its reprinting with additional material in his Contingencies of Reinforcement: A Theoretical Analysis (1969).
However:

. . .addition of instructions about the desired response results in rapid adoption of the response. . .but also may induce inappropriately high rates. . .More detailed instructions about reinforcing contingencies, as well as the response itself, typically produce response rates approximating the requirements of the reinforcement schedule. . .Finally. . .instructions about the reinforcement schedule may have effects overriding those of the schedule itself. Thus, instructions can induce behavior in the absence of reinforcement. . .and can produce behaviors more in accord with instructions than with actually scheduled reinforcement. (ibid.)

These authors conclude from their own study that the experimental analysis of behavior research program is not threatened by the recognition of instructed behavior:

. . .instructions represent an external, observable, determinant of behavior whose influences, although complex, can be investigated in a straightforward, objective manner. (p. 711)

In fact, the use of instructions may enhance behavioral research by making it possible to evoke behaviors that would otherwise not be brought into an experimental analysis. Indeed, this is entirely consistent with an extensional behavioral science research program. In the 1970s, explanations of the role of instructions in accounting for the animal–human discrepancies in schedule effects have been sought in, for instance, methodological procedures—such as the need to reinforce instructed behavior before the instructions will have an effect—and elements of the radical behaviorist canon such as learning histories—at least those acquired in the laboratory (Galizio, 1979).

The intervening years have strengthened many aspects of this early evaluation but have also raised the problem of subjects’ modifying instructions for themselves or even coming up with their own instructions and subsequently acting on the basis of these (at least partly-) self-formulated instructions rather than either the schedules or the experimenter’s instructions (for reviews see Baron & Galizio, 1983). One important strand of research is Catania’s which has established that the origins of verbal behavior influence its effect on nonverbal responding. Contingency-governed verbal instructions control verbal responding; rule- or instruction-governed verbal responses were inconsistent in their effects on nonverbal behavior (Catania, Matthews, & Shimoff, 1982). Lowe (1979) points out, however, that instructions do not indiscriminately lead to insensitive schedule performances and may even be essential to a subject’s performing schedule-programmed response patterns. It is the kind of instructions given that, above all, seems to influence sensitivity. Moreover, even contingency-shaped behavior, accompanied by a minimum of instruction, does not conform to the patterns of schedule-governed responding typical of animals.
The 1980s also saw experiments in which human performance on concurrent VI schedules was compared with that of animals, the latter having produced the phenomena of matching. The broad conclusion of this work was that

. . . under various conditions, the majority of [human] subjects performing on multiple concurrent variable-interval schedules showed gross departure from the matching relationship with forms of responding not previously encountered in animal studies of concurrent performance. As has been observed in previous studies of human performance on single schedules, there was considerable variability in responding, both within and between subjects. (Lowe, 1983, p. 76)

Self-Formulated Instructions

A common source of explanation of the observed differences in performance, both between humans and non-humans and among different human participants, has lain in attributing them to self-formulated instructions. Two separate questions arise: (a) What is the origin of these instructions? and (b) What is their consequence for the explanation of behavior? We should like first to explore in greater detail the answers to these questions provided by the behavior analysts whose experimental work led to the recognition of the discrepancy between animal and human behaviors, and then to examine the implications of their answers for radical behaviorism as a philosophy of psychology.

Verbal behavior and linguistic usage have long been suspected as the source of discrepancy. The agents of change may arise not only in the instructional behavior of the experimenter but equally in the capacity of experimental subjects to devise for themselves formulations of the contingencies in operation which guide their subsequent behavior. Nor is this at odds with radical behaviorist thought. As Lowe is at pains to point out, since at least 1945 Skinner argued for the inclusion of private events—thinking and feeling—in psychology and indeed defined radical behaviorism as uniquely incorporating such events. He insisted further that these were collateral responses, themselves the consequences of the extra-dermal contingencies that produced associated overt behaviors, rather than stimuli. The difference is that the later Skinner acknowledged that private events could have a causal effect on other behaviors (Skinner, 1988). It is presumably in this that other radical behaviorists have sometimes been reluctant to take the role of private events on board but Lowe and his colleagues have argued strongly for a resolution of the animal-human behavior schedule performance discrepancy that incorporates this level of analysis. Noting the tendency of human subjects on FI schedules to show either very high or very low rates of responding rather than the pause-respond pattern typical of animals, Harzem, Lowe, and Bradshaw (1978) examine three possible sources of discrepancy: response cost, conditioning history,

---

6 The term “extra-dermal” elicits the question whether there are “intra-dermal contingencies” and how they would be identified. This requires a longer response than is possible here. A starting point for answers, however, is Lee’s (1994) essay.
and verbal instructions. Response cost, in terms of either muscular effort required to execute a task or loss of points, does not account for inter-subject differences in response rates. Conditioning history undoubtedly results in some difference in response patterns: subjects whose previous low rate of responding on FI schedules has been differentially reinforced continue to respond at a low rate, while those whose behavior has been reinforced on a FR schedule respond at a steadily high rate. Conditioning history appears, moreover, to exert a stronger influence on current rate of responding for humans than for animals. Accounting for this as well as the inability of any schedule induced learning history to produce the post-reinforcement pause in human responding raise interesting theoretical questions for the experimental analysis of human behavior (Lowe, 1979). Finally, they note the tendency of instructions to alter response rates even when they are “contradicted” by the programmed rate of reinforcement (Harzem et al., 1978, pp. 406-407; see also Lowe, 1979). As Kaufman, Baron, and Kopp (1966) demonstrated, different types of instructions, ranging from minimal instruction through response-indicating instruction, VI instructions, and FI instructions, all produced different patterns of responding from both one another and from the programmed reinforcement rates. Harzem et al.’s (1978) own experiments indicate that learning history and verbal instruction can “greatly” affect human operant performance. Their conclusion is that

...in order to explain these phenomena, it is necessary to take into account the fact that the human subject in an operant experiment can verbally describe to himself the contingencies in operation, and that such description itself comes to act as a controlling stimulus. (pp. 420-421)

The discrepancy is also explained more elaborately in terms of human subjects’ behavior being “not under the control of experimental variables but controlled by self-produced cues, which vary from subject to subject” (Lowe, Harzem, & Hughes, 1978). These conclusions are supported by questionnaire evidence of the subjects’ self-rule formulation and counting procedures during the experiment, though some subjects relied on the provision of a digital clock rather than counting to determine their response rates. Lowe (1979) examines the role of self-instructions in the determination of human subjects’ response patterns by reviewing a number of experiments by other researchers. When subjects were interviewed after an experiment in which they received minimal instructions (Lippman & Meyer, 1967), some reported they thought that a number of responses was required to generate reinforcement; all of these conformed to the high response rate pattern in the experiment. The remainder reported that they thought reinforcers were made available on a temporal basis; all but one of these conformed to the low response rate pattern, the exception exhibiting the high rate. Lowe (1979, p. 169) summarizes:

The Lippman and Meyer study shows then a correlation not only between response pattern and externally-provided instructions but also a correlation
between responding and the subject’s own verbal formulation of the contingencies.

Similar correlations between an interval-based self-formulation of the contingencies and a low response pattern, and between a response-based self-formulation and a high response rate is reported by Leander, Lippman, and Meyer (1968). Other experimental evidence for subjects’ self-dialogue affecting their response patterns is provided by Laties and Weiss (1963), of whose results Lowe (1979, pp. 170-171) comments,

[It is difficult to explain the schedule performance of these subjects without taking into account the different self-instructional behaviors they adopted and the controlling effect exerted by such self-instructions.

From his own experiments, Lowe (1979) concludes that “what the human subject says to himself influences his behavior on schedules of reinforcement” (p. 185). The question that naturally arises next is where the verbal formulations come from that influence behavior, and why some formulations are interval-based while others are response-based. Several sources of the formulations are reviewed by Lowe (1979).

1. Learning History. An obvious source of such formulation is the kind of schedule on which the individual is responding: experience, current or prior, of FR schedules is likely to induce a response-based mode of behaving (i.e., high response rate) and to inspire reports of response-related sources of reinforcement. DRL schedules are more likely to induce interval-based modes of responding, low rates, and to inspire reports of temporally-based access to reinforcement. The subject on an FI schedule will receive reinforcement whether he adopts the low-rate mode or the high-rate mode; therefore, a subject who is switched from an FR to an FI schedule is likely to continue to responding at a high rate, showing the “insensitivity” to the schedule that is characteristic of human performance. The contingencies will have changed but not the subject’s formulation of them and this formulation controls behavior even though some reinforcement is thereby sacrificed. The naïve human subject encountering FI schedules for the first time in the laboratory may draw upon a more general learning history involving response generated reward and, depending on the precise nature of that history, exhibit either a low- or a high-rate mode of responding.

2. Experimenter Instructions. As noted, such instructions can result in either a low- or high-response mode, presumably, Lowe argues, by changing what the subject “says to himself” (Lowe, 1979, p. 187). Some researchers have suggested that not providing instructions might prevent this but

...much of the evidence presented here indicates that when contingencies are not made explicit to the subject he produces his own account of the relationship between his behaviour and the reinforcer, which may or may not coincide with the contingency in operation but which nevertheless continues to control behaviour. (p. 187)
Applied behavior analysis studies corroborate the argument that behavioral change requires description of the contingencies. If none is provided, individuals will produce their own (e.g., Lowe, Horne, & Higson, 1987; Meichenbaum, 1977).

The explicit acknowledgement of this capacity of the individual to produce his own rules, even though they run contrary to those implicit in instructions provided by an experimenter, distinguishes the approach to human behavior put forward by Lowe. The ability to “self-tact” is not new to behavior analysis, having been described by Skinner in Verbal Behavior (1957, p. 138), but as we shall argue it has far-reaching theoretical implications for radical behaviorism. Even more interesting in this regard is the capacity to employ their descriptions of the contingencies, and even erroneous descriptions of the contingencies, in devising rules that can and do govern their behavior and thereby render humans’ operant performance unique in the animal kingdom. Even humans who lack language do not exhibit operant performances that are peculiar to verbal humans: pre-verbal children, for instance, show similar response patterns to those produced by animals on programmed schedules (Lowe, 1979).

**Covert Explanation: Private Events as Causal**

The search for the causes of behavior has moved on now from the contingencies themselves to the verbal behavior of others to something that radical behaviorism has never entirely denied but certainly has de-emphasized: the creative verbal behavior of the individual. While the existence and scientific importance of private events has always been part of Skinner’s system, they have usually been denied causal significance in a science of behavior. Lowe’s solution is to make greater use of the private events that characterize radical behaviorism. He employs the writings of Skinner to the effect that covert behavior enters into the causation of overt behavior. Many radical behaviorists have shied away from this move because it is equally the case that Skinner has also written at some length of private events as collateral responses rather than causative stimuli; moreover, private events cannot by their very nature enter into an experimental analysis. In appealing to them for explanatory completeness, we are inevitably in the realm of interpretation. We shall argue here, first, that many verbalized private events are inherently intentional, and, second, that instructions cannot be understood other than in terms of intentional idioms.

The theoretical import of what is being said here derives from the acceptance that a verbal formulation of the contingencies is capable of influencing behavior and that radical behaviorism must incorporate this factor in its causal account of

---

7 While embracing the view that private events should be part of the causal fabric of radical behaviorism, I have endeavored to suggest the means by which such interpretation might be constructed in line with the usual canons of scientific enquiry (Foxall, 1990/2004, 1994, 1996, 2004, 2005, 2007a, 2007b). This endeavor has taken the form first of a model of consumer choice and second of a broader integrated psychology, intentional behaviorism, which informs the present examination of the role of private events in radical behaviorist interpretation.
behavior. It makes no difference if it is claimed that the formulations are themselves subject, at least in part, to the formulator’s learning history. It is the capacity to carry out the procedures involved in the formulation of the contingencies and to act upon the formulation that is central. There is no suggestion that the act of formulation is spontaneous, uncaused. Interesting ramifications, however, are, first, that such a formulation requires the capacity to self-describe and compare contingencies; second, that such formulations cannot be incorporated into an experimental analysis; and, third, that acting upon a formulation also requires the capacity to represent the formulated contingencies and to transduce this representation into motor activity. We need at least a competence theory which will involve intentionality; many would argue that we need a performance theory in which the cognitive operations involved in these procedures are examined\textsuperscript{8}. Is not “self-tacting” inherently intentional?

**Intentional Consequences**

An account in these terms cannot be avoided if we are to stay within the sphere of radical rather than methodological behaviorism. There need be no objection to the explanation given per se. However, it is evident that the language used here has taken us from the explanatory basis of radical behaviorism into a quite different realm of discourse and behavioral explanation, that of intentionality. The essence of Brentano’s nineteenth century rediscovery of the work on intentionality of the medieval scholastics is that terms that apparently relate to the mental produce sentences of a different kind from those that deal only in the physical (Brentano, 1874). The mental sentences are those that contain verbs of propositional attitude such as *wants*, *needs*, *thinks*, *feels*, and *intends*. This stems from the observation that such words, known as the “attitudes,” refer to operations that are always *about* something other than themselves: *aboutness*, we noted briefly above, is the heart of the philosopher’s understanding and definition of intentionality. The first consequence is that, following the use of these “attitudes,” codesignative propositions cannot be substituted for one another: it might be the case that “John knows that this is the flag of the United States,” but we are not justified in therefore saying that “John knows that this is the Stars and Stripes” since John may not know that the two terms refer to the same item. The extensional sentences of science by contrast permit such substitution: we can say not only “That flag is the flag of the United States” and “That flag is the Stars and Stripes,” exchanging codesignatives without having to be concerned who knows what. Second, the object of such sentences need not actually exist in the world: we can believe this or that about golden mountains, for instance. But the objects of extensional sentences must exist: if we say that “John is traveling to Edinburgh by car” there has to be a place called Edinburgh and a car to get there in. Third, we cannot translate intentional sentences into extensional without adding meaning: the

\textsuperscript{8} Dennett (1978) explains competence and performance theories. Note that Staddon (2001a, Chapter 6) presents a behaviorist critique of the competence–performance distinction.
two types of expression are not equivalent. “John said that he was going to Edinburgh” might be translated as “John said ‘I am going to Edinburgh’” or “John said, ‘I am off to the capital’” or “John said ‘I am leaving now for Old Reekie’” but in each case these add information and, hence, meaning.

Now, many private events—thoughts certainly, and some feelings—are intentional, always about something other than themselves. This is true from the first-personal subjective view, what Russell (1912) called knowledge by acquaintance, and from the first-personal or third-personal objective viewpoint of propositional attitudes (Russell’s knowledge by description). At the level of first-personal experience these events, qualia, are about something other than themselves: we think about, feel something else even our own bodies as Skinner says10, all of the other behaviors normally referred to as mental or cognitive—perceiving, remembering, hoping, fearing, and so on—all share the aboutness that is the essence of the philosopher’s definition of intentionality. Any attempt at expressing these experiences in language, be it by the person whose experience it is or by the observer, must employ the language of intentionality. The resulting sentences different from those on which extensional behavioral science is built: they do not admit the substitution of coextensives, they permit the phenomenon of intentional inexistence, and they are not reducible to extensional sentences. But this does not prevent their being conceptualized as discriminative stimuli and thereby taking their place in an extensional behavioral science. So, while we have raised the possibility that behaviorists are using the language of an antithetical mode of explanation without realizing it, we have not shown that behaviorism must adopt intentional explanation in its quest to predict and control11.

However, there is one area where it is impossible to avoid intentional language and thus intentional explanation: our description of the self-rule formulation process cannot but be expressed in terms of the individual’s comparing, planning, determining, and so on. At this point the radical behaviorist has switched from a purely extensional description of behavior to one based on intentional idioms; nor can these be avoided by an attempt to translate the intentional sentences into extensional accounts.

Discussion

It may not be immediately apparent how far this is explanation is removed from radical behaviorism as generally conceived. Yet, it strikes at the heart of radical behaviorist explanation in two ways.

9 See McGinn (1991, 2004). Does Skinner’s insistence that there is no subjective/objective distinction mean that knowledge by description is de facto omitted from a behavioral analysis?

10 Though he, like anyone else making this claim, needs to explain who or what it is that does the feeling!

11 Note that we are speaking here about differences between types of sentence, not between physical and mental realities (Chisholm, 1957; Dennett, 1969).
First, it admits the causal efficacy of private events. Skinner’s initial formulation of radical behaviorism included private events as collateral responses but not as stimuli, though he later acknowledged that covert verbal discriminative stimuli might function as “non-initiating” causes as long as it was recognized that the inaugural causes of behavior (the independent variables that shaped and maintained behavior) lay in the external environment (Skinner, 1988). The present formulation however attaches greater causal significance to private events since the thinking and feeling with which we are concerned cannot be easily traced to a learning history that establishes them as discriminative stimuli: they are after all initiated by the individual actor. Yet, because private events are not amenable to experimentation, we cannot know that they are causal in terms of the normal cannons of scientific judgment: just citing Skinner to this effect is not enough for a scientific analysis. It is simply to believe that private events are, faute de mieux, causal entities. What Lowe is proposing is that it is legitimate to delve into the efficient mechanism of causation that would have produced a particular pattern of observed operant behavior because this is not (entirely) apparent from our knowledge of final causes. But once we have sought efficient causation in private events, other putative mechanisms, such as those that bring about behavior as a result of thinking and feeling are open to scrutiny via further theoretical analysis. Even if private events are accepted as non-initiating or proximal causes while the ultimate or real causes of behavior lie in the external environment, we have to account for how private events can eventuate in overt behavior. In the absence of an experimental analysis, we cannot demonstrate that private events are discriminative stimuli or motivating operations that enter into a three- or n-term contingency to generate overt responding. We are in the realm of plausible surmise and, while such exploration is to be welcomed in a behavioral science that has tended toward dogma rather than the establishment of a sound methodology of interpretation, we have to acknowledge that once we have entered it there may be an infinite regress in the search for private causation. At the same time, we have to accept that such knowledge is tentative. Lee (1988) refers to interpretations as “hypotheses” but this is misleading since they are never testable in the manner of extensional behavioral science. We must find other ways of rendering our interpretations plausible and acceptable. The canons of procedure and judgment laid down by extensional science do not extend to the realm we have entered. In the absence of a systematic methodology of interpretation, radical behaviorist explanation relies upon speculation, particularly in the sphere of ascribing intentionality in order to provide an account of non-contingent behavior. The danger is that radical behaviorism will commit the same “offences” as those structural psychologies that invent causal entities at will.

12 The distinction is due to Aristotle. It is put most succinctly in the context of behavioral science by Rachlin (1994).
13 The establishment of satisfactory methodological grounds for the pursuit and evaluation of radical behaviorist interpretation, particularly the role of intentionality in such an
Second, this analysis requires extensive reconceptualization of radical behaviorism insofar as it involves intentional explanation. The admission of individuals’ deliberating, formulating, and deciding into an account of operant behavior is not intentional only because of these words’ status as propositional attitudes. It rests also on an argument advanced by Charles Taylor (1964) to the effect that the fundamental building blocks of behaviorist explanation, stimulus and response, are inevitably intentional. In order that the explanation of observed patterns of human schedule behavior can incorporate causal private events, the latter must be clearly specifiable, capable of definition in ways that allow them to be precisely delineated and demarcated from other events that can be shown to have no causal efficacy. An extensional behavioral science requires this. This is a very similar argument to that advanced by Taylor in his examination of the status of the term “stimulus” in operant studies of animal behavior. The problem is that although knowledge of stimulus control permits us to predict (and control) a series of responses, it can never explain why similar stimuli evoke similar responses or answer the question what is this similarity? Why do stimulus and response generalization occur? Indeed, to explain the continuity of behavior on the basis of stimulus control we would need (a) scientific definitions and measures of stimulus and response, so that they can be unambiguously recognized on subsequent occasions, (b) a complete learning history. We have neither. A rat’s jumping towards the white card and this being rewarded may not lead to reinforcement if the rat is not “paying attention” to what is going on (Taylor, 1964, pp. 124-125):

Hence on the cognitive view it matters what the rat is doing, that is, what action he is performing, and thus what intentional description the action has for him, whether “jumping right” or “jumping to white”, whereas on the S–R view, the response is not an action, the intentional description is irrelevant, and it matters only what description the card actually bears to which the rat jumped. (p. 125)

Taylor argues further that a stimulus could be any of the following: 1) a card or marks on card (the “distant stimulus”); 2) light waves from a distant stimulus that impinge on the retina; 3) proximal stimulus: pattern of excitation on receptors such as the retina; or 4) afferent impulses from reception to the brain. Only be taking into consideration the selective attention of the animal can a causal account be given and that is an account of a different order from the kind that radical behaviorism expects. Now an exactly similar problem of stimulus unspecificity arises in the case of interpretations of the private events that may

account, is the subject of Foxall (2004), and its application to the sphere of economic behavior has been undertaken in Foxall (2005, 2007c).

14 This argument is not met by the claim that it would be possible to ascertain through an experimental analysis the precise part of the card that was acting as the stimulus. Taylor’s point is that the subjective intentionality of the rat enters into a complex understanding or explanation of the behavior of interest. The objective viewpoint of the experimenter cannot provide this understanding for all its capacity to enhance prediction and control.

98
control overt responding. How are we to know how to label the private events that enter the causal chain? To isolate those that have a causal effect from those that might accompany them but have none? How do we know what the individual is responding to in the external stimulus field and, more importantly, among the possible private events that could influence his behavior? Taylor’s construal of animal behavior as explicable only in terms of selective perception is entirely homologous with Lowe’s construal of human behavior governed by private events. For the behavior of the human experimental subject who figures out the contingencies for himself and formulates a rule that governs subsequent schedule responding is exactly like that of the animal confronted with the card in that it rests on a unique personal interpretation of the contingencies that can influence behavior even when it does not correspond with reality as defined in the contingencies.

Particularly if rules are attributed to others in the course of radical behaviorist interpretation, their expression is intentional. Statements that take the form, “He is thinking that if he does A then B will follow,” which imply that his behavior is explained by his accepting/acting upon this track or that ply, consist in beliefs rather than rules derived from a learning history or either contingency shaping or instruction or both. We cannot conclude other than that there are many unknowns in this situation that we can only guess at; and that our best efforts at overcoming this problem involve the use of intentional language; this leaves us open to the charge that we are using a different form of explanation from that which radical behaviorism prescribes.

However, in accounting for the schedule-insensitive behavior of the human respondent in terms of his having formulated non-corresponding private events that govern behavior in the experimental setting, the language of rule-governed behavior is no longer adequate. We cannot say that the person has formulated a rule for clearly this is not the case: a rule is by definition derived form the contingencies, whereas the formulation that the subject is following is not: it is an invention, albeit possibly influenced by some kind of learning history that has itself been misconstrued. Surely, it would be safer to say that this person’s behavior is the result of his believing such and such that than that it is rule-governed. To say that it is rule-governed implies that we can point to the contingencies that have been formulated into a rule, the elements of the n-term contingency that can have been instructed or extracted. But, in this instance, we cannot. We have stepped into territory in which the falsification of our explanation is impossible. The construal of the contingencies that a rule represents by definition cannot be checked out by reference to any extra-personal reality. The person’s behavior cannot by definition as contingency-insensitive have ever been a response to known contingencies; so how can it be construed now as subject to a rule that embodies or reflects those contingencies? One response is that it must reflect contingencies to which the individual has been exposed at some time. . .this is wholly inadequate in behavior analysis.

There is undeniably a place for rules in an extensional behavioral science: the sounds uttered or sensations thought can be classed as physical discriminative stimuli or motivating operations that enter into the n-term contingency as would a
green light in an operant chamber. They are simply stimuli in the presence of which the individual performs particular responses that have previously been reinforced in the presence of those stimuli and they have come to exert some degree of controlling influence over the stimuli now; or they are antecedent stimuli that control the three-term contingency that follows them. We have no need to classify them in this case as intentional, though our researches will progress no further than the ability to predict and control behavior if we do. The experimental analysis of behavior can function perfectly well if these ontological and methodological assumptions are made. Beyond the pursuit of the limited goals of radical behaviorist inspired behavior analysis, it would then be feasible to leave talk of beliefs and desires—intentional talk—to others and get on with the job. Behavior analysis could proceed in the terminology of the extensional behavioral science that radical behaviorism enjoins and avoid notions of belief and desire entirely.

But there is a problem with this reasoning: we have no language in which to pursue the processes of calculation that does not involve intentionality. For, while we can portray the rule “You have to press the key a fixed number of times in order to earn the points” as a legitimate part of an extensionally described contingency, we cannot do the same for the rule “I need to press this key 17 times to get any points,” especially if the computer is programmed with an FR9 contingency between key presses and points delivery. Even if we could get at the self-rule that I have devised here (perhaps you will accept that this is the rule I have been following when I tell you), it cannot be said to enter into any causal contingency procedure in the way a discriminative stimulus or motivating operation would, for where is the training, the pairing of stimuli, and so on? The origin of the rule is my own deliberation and calculation; its guidance of my behavior is not the result of any training. The “rule” on which I acted was not contingency-shaped; it was instructed only to the extent that I calculated it should take the form it did. There is no way to describe this except intentionally. You can only say that I behaved this way because I believed that the contingencies were set up in the way described by my rule. My own misperception of what was going on when I was an experimental subject produced that rule, albeit aided and abetted by a learning history of rule following and perhaps of deviance. But all that, even if it were empirically available, would not be sufficient to explain my behavior on this occasion.

We have also to distinguish between an instructed rule, accurate or inaccurate, that guides and individual’s behavior, and an inaccurate rule that is the result of his own private configuration of the contingencies. The former may be said to bear some relation to the contingencies that have previously shaped the individual’s behavior in similar contexts; we may also conjecture that the person’s following the first kind of rule is partly or fully explained by a history of other-rule compliance in such contexts. An inaccurate rule is, by definition, attributable to neither of these sources. They are an influence on behavior that be neither traced to environmental contingencies nor plausibly interpreted as the result of prior verbal instructions (unless the experimenter has deliberately provided erroneous
INTENTIONAL CONSEQUENCES OF SELF-INSTRUCTION

information: in this case they might be attributable, though usually without any evidence whatever, of course, to a general tendency to follow the instructions of persons in authority).

We might propose a history of successfully following self-rules to account for this. But this fails for three reasons. First, it sounds implausible and suggests an infinite regress of putative self-rule formation and following, the unrestrained multiplication of explanatory fictions; second, it does not account for how, on this occasion, the individual misconstrued the contingencies without cognitive input (i.e., an ability to represent the contingencies privately and make erroneous calculations about the consequences they portend; such causal representation and calculation being the very hallmarks of cognition); third, it cannot be attributed in any case to a history of following inaccurate self-rules since the outcomes of doing so cannot by definition have been reinforcing.

Moreover, whatever the origin of the private behaviors that leads to the self-defeating schedule performance, such private behavior adds something that is apparent from its outputs (conclusions about how to behave) but which cannot be traced to stimulus inputs. This is a hallmark of the attribution of cognitive activity in humans and non-humans.

Conclusions

Empirical research conducted firmly within the radical behaviorist camp, and the significance placed by radical behaviorists themselves on the results of that research, indicate that radical behaviorism, having admitted the possibility of rule-governed behavior, can no longer maintain its exclusively extensional methodology. If radical behaviorism goes beyond the analysis of contingency-shaped response, then it is not only obliged to adopt intentional language: nor can it avoid cognitive explanation. It is specifically the possibility of explaining overt behavior in terms of private events, especially when these are in some degree self-generated, that prompts the need for further conceptual revision. In particular, having agreed that intentional ascription is an inescapable part of explaining human behavior, we face the question of how content is to be used legitimately: radical behaviorism has always and quite rightly railed against the fanciful adoption of intentional terms as readily-available “explanations.” Dennett’s (1969) answer—to ascribe intentionality only if it is consistent with naturally-selected afferent–efferent neuronal links—suggests a key. One derivative of this approach, intentional behaviorism\(^\text{15}\), retains Dennett’s insight but adds a behavioral criterion for content ascription: the patterns of molar behavior maintained by environmental

\(^{15}\) It is, of course, incumbent upon anyone proposing a new kind of behaviorism to examine critically its relationship with other schools of behaviorism. In the case of intentional behaviorism, this ongoing goal has been accomplished to some extent: Foxall (2004) examines the relationship between intentional behaviorism and radical behaviorism at some length, as well as comparing both with teleological behaviorism (Rachlin, 1994) and, briefly, to theoretical behaviorism (Staddon, 2001b). Further critical comparisons can be found in Foxall (2007a, 2007b, 2007c).
contingencies that behavior analysis investigates. In this formulation, intentional explanation at the level of the person is corroborated by two extensional sciences, neuroscience at the sub-personal level and behavioral science at the super-personal level.16

As a result, two kinds of conceptual restructuring are required. The first is the acknowledgement that intentional terms cannot be avoided in the explanation of human behavior and hence that extensional behavioral science, whilst an essential component of treating such behavior, is of itself inadequate to explain it in its entirety. Second, the proposal that human reasoning and the representation and private manipulation of contingencies occurs as individuals formulate self-rules that govern their behavior may require acknowledgement that cognitive processes occur and must enter into the explanation of some dimensions of human behavior.

Behavior analysts who resort to private events to account for behavior that cannot be explained in terms of observable contingencies are in effect using the intentional stance (Dennett, 1987): attempting to predict the behavior of the subject by attributing to him the beliefs he ought to have in order to account for his observed behavior. Moreover, they are using the methodology that Dennett (1991, 2005) later called heterophenomenology in order to reconstruct at a third-personal level the phenomenology that would explain that behavior. We have no objection to this, we have not argued against it, and we believe it to be an inevitable consequence of trying to explain rule-governed behavior in terms of private events. But it is a long way from the three-term contingency.

References


16 Its full implications for radical behaviorist interpretation remain beyond the scope of the present paper; see Foxall (2004).
INTENTIONAL CONSEQUENCES OF SELF-INSTRUCTION


ta, 114*, 107-116.


