INTENTIONAL BEHAVIORISM

Gordon R. Foxall
Cardiff University

ABSTRACT: Two of the leading contenders to explain behavior are radical behaviorism and intentionality: an account that seeks to confine itself to descriptions of response–environment correlations and one that employs the language of beliefs and desires to explicate its subject matter. While each claims an exclusive right to undertake this task, this paper argues that neither can be eliminated from a complete explanatory account of human behavior. The behavior analysis derived from radical behaviorism is generally sufficient for the prediction and control of behavior in the laboratory and its applications, but it fails to provide an explanation of behavior since it cannot deal with the personal level of explanation, the continuity of behavior, and the delimitation of behaviorist interpretations. Only the inclusion of intentional terms can achieve these ends. An intentional account cannot succeed, however, without the incorporation of a behavioral criterion for the ascription of intentional content based on the analysis of systematic environment–behavior relationships. This paper proposes an overarching philosophical framework for the analysis and interpretation of behavior that incorporates both radical behaviorism and intentional psychology in a model, "intentional behaviorism," that additionally links the explanation of behavior to neuroscience and evolutionary psychology. Finally, the paper proposes a link between the philosophical framework of intentional behaviorism and the world of empirical science by describing a tentative model of research, "super-personal cognitive psychology," that shows how the disparate elements previously discussed impinge upon psychological investigation.
Key words: radical behaviorism, intentional psychology, philosophy of science, explanation, Dennett, Skinner

> [O]ne cannot expect the question as to the scientific status of psychology to be settled by empirical research in psychology itself. To achieve this is rather an undertaking in epistemology. (Hempel, 1980, p. 16)[1]

[1] Hempel's paper was originally published in H. Feigl & W. Sellars (Eds.) (1949). *Readings in Philosophical Analysis* (pp. 373-384). New York: Appleton-Century-Crofts. Reference here is given to its republication in 1980 since this source contains revisions by the author and a brief account of how his thought changed since the initial publication.

Introduction

The more I see of behavior analysts' struggling to accommodate the thinking of other psychologists, usually cognitive in orientation, to the vocabulary of behavior theory, and thereby missing much of the point, including much of what is relevant to behaviorism, the more I wonder why we are not open to a broader philosophical framework that can cope with both kinds of utterance. The work of the famed behaviorist philosopher Quine (1960) should be sufficient to alert us to the impossibility of such translation. Moreover, numerous other philosophers such as Chisholm (1957), Dennett (1969), and Searle (1983) at least suggest to us the explanatory riches of a linguistic system that includes intentional idioms such as desires and beliefs, and a range of empirical analyses of behavior—among others, those of Bolles (1972), Bindra (1978), Toates (1986), and Dickinson (1997)— make clear the advantages of a psychological theory that recognizes the two. Some behavior analysts readily employ the language of intentionality, apparently oblivious of the extra-behaviorist avenues of explanation into which this necessarily leads them. But most of us seem to prefer a vocabulary that is limited to the descriptive level of the three-term contingency and thereby restricted in terms of the range of explanation open to us and the contribution that we, as behavior analysts, can make to the development of both empirical and theoretical psychology. I should like to explore a framework of conceptualization and analysis, "intentional behaviorism," that embraces the terminology of radical behaviorism and intentional psychology, not because such a synthesis is desirable on its own merits, but because each of the systems of explanation represented by these linguistic modes is necessary to the completion of the program of the other.

Intentional psychology, in which beliefs and desires assume a central explanatory role, provides the foundation of cognitive psychology and much of its social, organizational, educational, and economic applications. Its explanatory stance is exactly opposite to that of behaviorism, which has traditionally striven at all costs to avoid intentional terms like "believes" and "desires." However, despite the fact that neither program can succeed wholly without the other, there is much that the proponents of each of these approaches have misunderstood in the other's arguments.

Finding a resolution thus requires a thorough and critical examination of both intentionality as a means of explaining behavior and of radical behaviorism as a particular philosophy of psychology. Hence, this paper draws upon the contributions of two leading exponents of these respective systems: Daniel Dennett in the case of intentionality and Frederic Skinner for radical behaviorism. This is necessary in order to understand how their systems contrast with alternative approaches to intentionality and behaviorism, respectively, as well as to build a coherent, overarching framework for understanding behavior.

Dennett's program, since the appearance of his first book, *Content and Consciousness,* in 1969, has been concerned with the place of intentional idioms in the explanation of behavior and the neurological basis of consciousness—in other words, in the biological and philosophical underpinnings of cognitive psychology.

Central to both enterprises has been the distinction between a *sub-personal* and a *personal* level of analysis. This distinction has informed his work on the legitimate ascription of content to intentional systems and the delineation of human consciousness. There is, nevertheless, controversy among philosophers over the significance of Dennett's distinction between personal and sub-personal levels of explanation and the uses to which it may be put. The disagreement is occasioned in some degree by the different emphases Dennett himself has accorded the meanings and usages of these levels since he introduced the distinction and by the different criteria he has emphasized over the years as appropriate to justify the ascription of content. Difficulties include the number of intentional explanations suggested by Dennett's successive analyses, the relationships among them, and the legitimacy of ascribing content at more than one level. While Dennett's initial distinction apparently precluded the ascription of content at the sub-personal level, which was identified with neuroscientific theory and research, his later tendency casts the personal/sub-personal distinction as that between the whole and its parts, with the result that the personal level as a source of explanation in its own right has been relatively ignored. The later thinking that permitted the ascription of content to sub-personal components via the pragmatic use of the intentional stance helped blur the original distinction between explanatory levels.

The intentional stance is the philosophical position that any entity the behavior of which can be predicted by attributing to it beliefs and desires is an intentional system, and this lays open the possibility of ascribing intentional content not only at the personal level but at any level that facilitates prediction (Dennett, 1983, 1987). The intentional stance is better understood when contrasted with the two other stances Dennett introduced at the same time. The *design* stance is used to "make predictions solely from knowledge or assumptions about the system's functional design, irrespective of the physical constitution or condition of the innards of the particular object" (Dennett, 1978, p. 4). The information provided by this stance leads us to define what an object will do, what its function must minimally be, regardless of its form. From the *physical* stance we make predictions on the basis of the physical state or conditions of the system; it depends on knowledge we have in the form of laws of nature. Predicting that when the bough breaks the baby will fall involves using the physical stance, as does forecasting that the atmospheric conditions that are about to bring rain will also bring on my lumbago. Through the recognition that the best chess-playing computers now defy prediction by either of these stances, Dennett arrives at the third stance: the *intentional* stance. In using it, ". . .[O]ne assumes not only (1) that the machine will function as designed, but (2) that the design is optimal as well, that the computer will 'choose' the most rational move" (Dennett, 1978, p. 5). Note that rationality here means optimal design relative to a goal, and that prediction is relative to the nature and extent of the information the system has about the field of endeavor. "One predicts behavior. . .by ascribing to the system *the possession of certain information* and supposing it to be *directed by certain goals*, and then by working out the most reasonable or appropriate action on the basis of these ascriptions and suppositions. It is a small step to calling the

information possessed the computer's *beliefs*, its goals and subgoals its *desires*" (Dennett, 1978, p. 6).

Dennett's associated attempt to formulate the philosophical basis of "sub-personal cognitive psychology" in contradistinction to the intentional systems theory that inhered in the personal level of explanation has enabled his system of explanation to become increasingly elaborate, but it has also increased confusion.

The issue is clearly of direct concern to the quest for the philosophical foundations of social cognitive psychology inasmuch as processing accounts necessarily rest upon some construal of the nature of intentional ascription, upon the specification of cognitive performance by an underlying competence theory of intentionality and behavior (Bechtel, 1988). I shall argue, however, that the distinction is just as important an element in the foundations of behavioral psychology, for the personal/sub-personal distinction suggests a personal/super-personal distinction that involves behavioral science in a more complete psychological science.

In contrast to Dennett's intentional psychology, Skinner's radical behaviorism avoids reference to cognitive events and processes in its explanations of behavior, repudiating cognitivism as the "creationism of psychology" and proposing that whatever believing and desiring may also be, they are behaviors to be characterized and explained like any other behaviors, invisibility to third parties notwithstanding (Skinner, 1974). Behavior is a function of its environmental consequences, which reinforce (make it more probable in similar circumstances in the future) or punish (make behavior less probable). I argue that, while such contingencies may be valuable for the prediction and control of behavior, they are inadequate for an explanation of behavior that seeks to account for its continuity and its representation at the personal level, and that is constrained by the credible consequences of the action. Only intentionality can provide the theoretical structures required to accomplish an adequate explanation of behavior. I nevertheless propose that the intentional program itself cannot be completed without a formal and systematic understanding of the role of the environmental determinants of behavior as reflected in behavior analysis.

Although the admission of private events (thoughts and feelings) into the ontology of radical behaviorism (Skinner, 1945) is usually taken to demarcate it from other neo-behaviorisms such as those of Tolman and Hull, it is in fact the determination of radical behaviorists to avoid intentional language in their statements of behavioral causation that is the defining characteristic of radical behaviorism. Yet, despite radical behaviorism's effectiveness as a means of predicting behavior, the question arises as to whether it can provide a satisfactory explanation thereof without resort to intentional idioms.

Intentional and Extensional

Some words seem to "reach out" to things other than themselves. They "refer" to or are *about* something else. In using the word "desire," it is necessary to specify what it is that is desired; in speaking of "belief," it is similarly impossible

to proceed without alluding to what is believed. Belief and desire are known by philosophers as "attitudes," and their sense is completed by stating a proposition about what is believed or desired. Such propositions usually begin with "that." I do not simply desire: I desire *that* something or other be the case. I do not just believe: I believe *that* such and such is the case. Not all words are of this kind. We do not speak of "breathing that" or "bathing that." The first kind of expression is called intentional (from the Latin *intendere*). While intentionality's special properties were known to the scholastic philosophers of the Middle Ages, it came to prominence with Brentano's nineteenth-century insistence that intentionality must be the mark of the mental. Mental events were to be distinguished from physical things by virtue of the way they were spoken of. In addition to the "aboutness" of such locutions, Brentano drew attention to two linguistic characteristics that separated them from other words.

First, it may not be truthful to substitute two words or phrases that are equivalent one for the other in intentional sentences. "John believes that Satan roams the earth like a devouring lion" is an intentional sentence that may be true, but it might not be accurate to say "John believes that the Devil roams the earth like a devouring lion" since John may not be aware that, at least in the context of fairly orthodox Christian belief, Satan and the Devil are one and the same. Sentences of this kind are said to be referentially opaque by dint of the incapacity of the speaker to substitute codesignative terms and yet be sure of retaining the truth value of his or her words. This contrasts with the *extensional* language of science, in which the substitution of codesignatives leads to perfectly accurate statements (which are, accordingly, known as referentially transparent). One can alternate between "There is the Prime Minister of Great Britain" and "There is the First Lord of the Treasury" without losing the truth value of what one is maintaining. The Prime Minister and the First Lord are one and the same person; hence, the sentences have the same extension.

Second is what Brentano called "intentional inexistence": the things referred to in an intentional sentence do not necessarily exist. When I say "I believe that there are Hobbits at the bottom of my garden" I speak of imaginary creatures that have no existence other than fictional—but when I say "I am going to drive my car to Scotland," there has to be a car which I shall drive.

The modern emphasis among some philosophers of psychology is that intentionality is not the mark of the mental in any sense that would suggest a sharp ontological dichotomy, but a means of distinguishing alternative ways of speaking of or explaining the world: a source of alternative modes of explanation: alternative and quite distinct, because it is not possible to translate intentional sentences into extensional ones that carry precisely the same meaning. Philosophers from Chisholm (1957) and Dennett (1969), for instance, who are sympathetic to the idea of incorporating intentionality in behavioral explanations, to the behaviorist Quine (1960), who is not, agree on this. The everyday, matter-of-fact observation "She said that the train would be late" cannot be rendered with certainty as "She said 'The train will be late'" since, by specifying a form of words that she uttered, this adds information that the intentional sentence does not

contain. For the same reason, "She said 'The train has been delayed'" or "She said 'The train's engine needs to be replaced'" or "She said 'The driver is drunk again'" are not equivalent to the intentionally-expressed sentence. Each kind of statement belongs to a distinct mode of explanation: one leads irrevocably to cognitive psychology; the other, to behaviorism.

Radical Behaviorism

Radical behaviorism as a philosophy of psychology is strictly extensional: it strives to account for its subject matter, behavior, in sentences that are referentially transparent, in which codesignatives are substitutable because they have the same extension. It is thus distinguished from cognitivism by its rigorous avoidance of intentional language, and from both cognitivism and other neo-behaviorisms by its inclusion of thinking and feeling ("private events") as phenomena that require explanation on the same terms as public responding. Its focus is the prediction and control of behavior by reference to its environmental consequences and the antecedent stimuli that set the scene for reinforcement or punishment. In its adherence to Machian positivism, it holds that when the environmental stimuli that control behavior have been identified, the behavior has been explained. The truth criterion it applies to this endeavor is pragmatism rather than realism. The scientific arm of this philosophy, behavior analysis, seeks the prediction and control of behavior in the environmental–behavioral contingencies which, in their familiar "three-term" construal, propose that S^D: R → S^R where S^D is a cue or *discriminative stimulus*, R is an *operant class*, and S^R is a *reinforcing stimulus*. The discriminative stimulus (S^D) sets the occasion (:) for (but does not elicit as does the unconditioned stimulus of classical conditioning) an operant class (R) which produces (→) a reinforcing consequence (S^R), which, via feedback, makes the future enactment of this operant class in similar circumstances more probable (Moore, 1999; see also Staddon & Cerutti, 2003). The behavior in question is *operant* behavior, which, by operating on the environment, brings about consequences that control its future rate of emission.

Each element of the three-term (or, in general, *n*-term) contingency is described in extensional language: its operation is not dependent upon wants or beliefs, desires or intentions (Smith, 1994). It describes both contingency-shaped and rule-governed behaviors in terms of "a system of functional relationships between the organism and the environment" (Smith, 1994, pp. 127-128). Hence, an *operant response* "is not simply a response that the organism thinks will have a certain effect, it does have that effect." Further, a *reinforcer* "is not simply a stimulus that the organism desires to occur. It is a stimulus that will alter the rate of behavior upon which its occurrence is contingent." And a *discriminative stimulus* "is not simply a stimulus that has been correlated with a certain contingency in the organism's experience. It is one that successfully alters the organism's operant behavior with respect to that contingency." Descriptions of contingent behavior do not take propositions as their object; rather, their object is relationships between an organism's behavior, its environmental consequences, and the elements that set the

occasion for those contingent consequences. So behavior analysis does not attribute propositional content to any of the elements of the three-term contingency. "Instead of accepting a proposition as its object, the concept of reinforcement accepts an event or a state of affairs—such as access to pellets—as its object" (Smith, 1994, p. 128). Mentalistic description: "The animal desires that a pellet should become available." The behavior analytic description is not "The animal's lever presses are reinforced that a pellet becomes available"; it is "The animal's lever presses are reinforced by access to pellets." A discriminative stimulus would not be described as a signal *that* something will happen, but simply that a contingency exists. "It attributes an effect to the stimulus, but not a content." Whereas the substitutability of identicals fails in mentalistic statements (such statements are said to be logically opaque), behavioral categories are logically transparent, suggesting that "behavioral categories are not a subspecies of mentalistic categories" (Smith, 1994, p. 129).

Neither is the proposition that "reinforcer" merely denotes "desire" feasible; desires are not equivalent to reinforcers, or reinforcers to desires. Common-sense notions imply that if a stimulus is (positively) reinforcing it is desired, and if it is desired it is because it is a (positive) reinforcer—but in fact neither notion holds. Objects of desire may not be attainable (e.g., the fountain of youth, perpetual motion) and so cannot be (linked to) reinforcers. Nor are reinforcers necessarily desired: given appropriate histories, responding under fixed and variable-interval schedules can be maintained by the *delivery* of electric shock, at least with monkeys and cats (see, e.g., Morse & Kelleher, 1977). The shocks are avoidable simply by not responding, but are not avoided. Noxious shocks (which otherwise can serve as negative reinforcers) cannot be "desired," yet under certain conditions they can maintain behavior as positive reinforcers.

Nor do functional units of the speaker's verbal behavior, such as mands and tacts (Skinner, 1957), have propositional content. They are simply statements of contingencies that account for an individual's behavior in the absence of his or her direct exposure to those contingencies. A mand is "a verbal response that specifies its reinforcer" (Catania, 1992, p. 382): for example, "Give me a drink" plus the unspoken, "You owe me a favor" or "Else I shall ignore your requests in future." Even if this is expressed as "I desire that you give me a drink. . .", it is actually no more than a description of contingencies. A tact is "a verbal discriminative response. . .in the presence of or shortly after a stimulus" (Catania, 1992 p. 399): "Here is the bank." Even if this were expressed as, "I want you to see the bank," its function would be confined to establishing the stimulus control of the word "bank," as when the listener replies, "Oh, yes, the bank." More technically, the *mand* denotes the consequences contingent upon following the instructions of the speaker or of imitating his or her example. Much advertising consists of mands— "Buy three and get one free!" "Don't forget the fruit gums, mum"—which indicate contingencies under the control of the speaker. *Tacts* present a con*tact* with part of the environment and, depending on learning history, a potential for behavior on the part of the recipient. A trademark or logo may be followed by making a purchase or entering a store. The definitive source is Skinner's *Verbal Behavior* (1957).

The functional units of the listener's verbal behavior, as proposed by Zettle & Hayes (1982), similarly describe contingencies rather than express propositional content. Pliance, for instance, is the behavior of the listener who complies with a verbal request or instruction: hence, "Pliance is rule-governed behavior under the control of apparent socially mediated consequences for a correspondence between the rule and relevant behavior" (Hayes, Zettle, & Rosenfarb, 1989, p. 201). Pliance is thus simply the behavior involved in responding positively to a mand. Tracking is "rule-governed behavior under the control of the apparent correspondence between the rule and the way the world is arranged" (Hayes et. al., 1989, p. 206). It involves tracking the physical environment, as when following instructions on how to get to the supermarket. Once again, its form—for example, "Turn left at the traffic light" plus the unspoken "and you'll get to Sainsbury's"—is a basic description of contingencies rather than an expression of propositional attitudes. Precisely as Smith has concluded with respect to contingency-shaped behavior, we may conclude with respect to rule-governance that "beliefs and desires have propositional content. . . .Designations of discriminative stimuli and reinforcing stimuli, by contrast, do not accept *that*-clauses" (Smith, 1994, p, 128). A third functional unit of listener behavior has no corresponding unit for the speaker: the *augmental* (Zettle & Hayes, 1982) is a highly motivating rule that states emphatically how a particular behavior will be reinforced or avoid punishment: "Just one more box top and I can claim my free watch!"

The private events that distinguish radical behaviorism are not "cognitive" or "mental" rather than material or physical. They are essentially private, collateral responses under the influence of the same environmental stimuli that control overt—or, better, public—responding. As such, their ontological status is fixed by their place in the three-term contingency; they are responses in need of operant explanation by means of an account that causally links them with antecedent and reinforcing stimuli occurring in the extra-personal environment rather than discriminative or reinforcing stimuli, which are capable of determining the frequency of a response. They are dependent variables.

Radical behaviorism explains verbal behavior in similar terms to nonverbal behavior: that of the speaker as a series of functionally defined speech (and quasi-speech) units—tacts, mands, autoclitics, echoics, intraverbals; that of the listener as a series of functionally defined verbal units that prescribe the consequences of rule-following—tracks, plys, augmentals.

Behavior analysis seeks to proceed extensionally, that is, in verbal behavior that avoids propositional content, describing its observation in language that is referentially transparent. It has three components or modes: (1) the experimental analysis of behavior, which is a laboratory-based investigation, (2) applied behavior analysis devoted to interventions to treat behavior dysfunctions, instructional design and execution, organizational behavior management, etc., and (3) radical behaviorist interpretation using the principles of behavior gained in basic and applied analysis to provide an account in operant-contingency terms of the complex behaviors that are not amenable to direct experimental examination. Radical behaviorist interpretation frequently involves the use of mediating events,

something ostensibly ruled out by Skinner's avoidance of "theoretical terms" but which appears necessary at this level of explanation. However, these mediating events are not intentionalistic; they remain part of an extensional account whose explanatory terms are extrapolated, for example, from the experimental to the non-experimental sphere.

Radical behaviorist explanation thus proceeds on the basis of the *contextual stance* (Foxall, 1999), which states that behavior is predictable insofar as it is assumed to be environmentally determined; specifically, insofar as it is under the control of a *learning history* that represents the reinforcing and punishing consequences of similar behavior previously enacted in settings similar to that currently encountered. The contextual stance thus portrays behavior as taking place at the temporal and spatial intersection defined by learning history and behavior setting. It is this intersection that defines the situation.

Three Imperatives of Intentionality

While there is no doubting the capacity of behavior analysis within the framework of radical behaviorism to predict and control behavior, in the operant laboratory (as well as in successful applications) there is a need for further conceptualization if we wish to account more fully for certain aspects of behavior. Explanation of the extensional kind is optional for behavior analysts, who may wish to remain within the philosophy of science set by Machian positivism as, did Skinner (Mach, 1896/1959, 1905/1976; Smith, 1986)—but there is no compelling reason to confine inquiry to this extensional level of analysis. In seeking to extend the conceptual framework here, I am concerned with methodology—with instances in which it is impossible to proceed with inquiry in the absence of intentional language—rather than with ontological questions. I should like to pursue three areas in which I believe explanation that goes beyond the *n*-term contingency can yield answers to questions that would be asked as a matter of course in most scientific endeavors but which have not usually found a place within radical behaviorism. These areas concern the treatment of the personal level of analysis, accounting for the continuity of behavior, and delimiting behavioral interpretations of behavior by delineating the scope of behavioral consequences that can be called upon to provide a causal explanation thereof.

The Personal Level

The personal level of analysis is central to Dennett's earliest work on intentionality. The personal level of explanation is that of "people and their sensations and activities" rather than that of "brains and events in the nervous system" (Dennett, 1969, p. 93). The latter belong to the sub-personal level, at which an extensional science such as physiology (neuroscience) operates, its mechanistic explanations inappropriate to so-called mental entities such as pain which occur and can be understood only at the personal level. The personal level is that at which the organism as a whole can be said to act. As Ryle and Wittgenstein

have pointed out, it is a stage of explanation that is quickly exhausted because so little can be said at this level. Of his pain, the bearer can say little more than that it hurts. In Dennett's system, as we shall see, this is the level at which beliefs, desires, and other intentional idioms are ascribed, but for now we are concerned only with the personal level as an analytical tool in extensional behavioral science and its implications for the explanation of behavior.

The personal level has two aspects: a first-personal perspective (that from which I actually feel my pain as an inner-body experience) and a third-personal viewpoint (that in which I attribute pain to another person who is sobbing and holding her head as well as using the word "migraine" a lot). The acceptance of these "subjective" and "objective" understandings of the personal level does not divide cleanly along behaviorist/non-behaviorist lines. Skinner's analysis of private events can be read as embracing both at one time or another. Dennett's cognitive approach concentrates on the objective, third-personal level, which he associates unremittingly with a scientific standpoint, while Schnaitter's (1999) behaviorist view is ready to endorse the first-personal. Others, such as Searle (1983), fully accept the necessity of speaking in terms of both the first- and the third-personal, and that is the approach that I take.[2]

The difficulty for radical behaviorism—or any other brand of extensional behaviorism—is that it deals inadequately with both first- and third-personal aspects of the personal level, largely because it confuses them. First note that in the case of the first-personal or subjective level of personhood, radical behaviorism simply has no means of accounting for some behaviors without resorting to intentional language. This stems from the irreducibility of intentional language to extensional language and is illustrated by the following examples of people acting contrary to their desires, beliefs, and expectations in ways that cannot be entirely captured in a purely extensional description. Take, for instance, the couple who found themselves married because they went through the motions of a Jewish wedding ceremony, they with all the other participants thinking that they were engaged in an elaborate joke, only to discover that they were, in fact, married. No one intended this outcome; one member of the couple fully intended to marry someone else. Another example concerns the Muslim acting with his real-life wife in a television production who, having followed the script to the letter, found himself divorced from both his screen wife and his actual spouse, unable to live with her on pain of being found guilty of adultery. This, again, was contrary to the expectations the entire cast and production team held about the situation (both examples are taken from Juarrero, 1999). The point is not that a radical behaviorist interpretation of these behaviors is impossible, or even whether they are actual or

---

[2] I have defended elsewhere (e.g., Foxall, 2007a) the incorporation of non-causal subjective experience into the framework of exposition I propose here. Such a view does not, of course, form part of Dennett's system of intentional ascription, but this does not affect the current argument since it is sufficient for present purposes to confine the personal level, as Dennett does, to a third-personal account. For the sake of completeness, however, I note that my view of this dichotomy is closer to those of McGinn (1991, 2004) and Strawson (1994)—and for that matter, Skinner (1945)—than that of Dennett (1969).

anecdotal, but that such an interpretation can never capture the entire behavior in question without resorting to intentional idioms (i.e., without deviating from its commitment to extensional behavioral science).

So how does it cope? Skinner's approach to interpretation is to seek the explanation of an individual's current behavior in his or her history of reinforcement and punishment (i.e., learning history). Despite the way in which the three-term contingency is usually symbolized as showing the factors that cause a response as the consequences that necessarily follow it, Skinner does not try to explain behavior by reference to future events. He avoids teleology by explaining current behavior in terms of the consequences that have followed similar responding in the past. Hence, when we see someone rummaging about among the objects on her desk, we infer that she is looking for her glasses. But the information available to us to make sense of her behavior is identical to the information she has to do the same. All she can say in explanation is that she has found her glasses in the past when she has engaged in behavior of this kind. The behaviorist strategy of "discovering" a learning history in order to interpret complex behavior evidently accords with the philosophy of behaviorist explanation (Baum & Heath, 1982). Although it eschews the mentalistic fictions Skinner so strongly repudiated, it nevertheless extends the analysis of human behavior beyond the confines of a scientific enquiry.

Very rarely, if at all, do we base statements about our emotions, say, on the kind of observation of ourselves that a third person would make. A person does not come to understand that he is nervous because he sees his hands shaking and hears his voice quavering. He does not come to conclude that he is nervous on the basis of evidence of this kind any more than his saying he has a headache depends on his prior observation of his flushed features, his holding his temples, and his having taken aspirin. As Malcolm (1977, p. 97) says, "If someone were to say, *on that basis*, that he has a headache, either he would be joking or else he would not understand how the words are used. The same is true of a first-person perception sentence, such as 'I see a black dog'." He argues further that behaviorists have erred by assuming that a psychological sentence expressed in first-personal terms is identical in content and method of verification to the corresponding third-personal sentence. We verify that another person is angry by the way the veins stand out on her neck, by the redness of her face, and by her shouting. But we do not verify our own anger in this way. We do not, as a rule, attempt to verify it at all. Verification is simply not a concept or operation that applies to many first-person psychological reports (those that are not founded on observation). An individual's statement of purpose or intention belongs in a different class from one made by someone else on the basis of observing that individual. If we see someone turning out his pockets and recall that on previous occasions he has done this before producing his car keys from one of them we can reasonably conclude that he is looking for this car keys this time too—but it would be odd indeed if he himself were to work out what he was doing by observing that he was emptying his pockets as he had done in the past when looking for his car keys. If he announced that he must be looking for his car keys at present because he was

doing what he had done in the past when finding them had eventuated, we should think him most odd and crazy, to be treated in future with circumspection.

Malcolm (1977, p. 99) also draws attention to speech acts such as "I was about to go home," which for Skinner present the problem that it "describes a state of affairs which appear to be accessible only to the speaker. How can the verbal community establish responses of this sort?" (Skinner, 1953, p. 262). Skinner's explanation is that as the speaker has previously behaved publicly, private stimuli have become associated with the public manifestations: "Later when these private stimuli occur alone the individual may respond to them. 'I was on the point of going home' may be regarded as the equivalent of 'I observed events in myself which characteristically precede or accompany my going home.' What these events are such explanation does not say" (Skinner, 1953, p. 262). Malcolm comments, "For Skinner 'private stimuli' would mean of course physical events within the individual's skin. The fact that Skinner regards this hypothesis as a possible explanation of the utterances, even though he does not know what the private stimuli would be, shows how unquestioningly he assumes that such a remark as 'I am on the point of going home' must be based on the observation of something" (Malcolm, 1977, p. 99). But the statement "I am on the point of going home" is not a prediction based on the observation of anything: "The announcement 'I am about to go home' is normally an announcement of intention. Announcements of intention are not based on the observation of either internal or external variables. . ." (Malcolm, 1977, p. 99).

Statements of intention are undoubtedly related to external events, and someone who said he was about to go home would normally have a reason for doing so, for example, that it was time for dinner. But this does not mean that going home or making the utterance is under the "control" (in Skinner's sense) of dinner time. In Skinner's technical sense of control, y is under the control of x "if and only if x and y are connected by some *functional relationship*," and if control is given this sense then neither intentions nor statements of intention are "controlled" by anything (Malcolm, 1977, p. 100). On the one hand is the claim of some behaviorists that "psychological" language (that which deals with so-called mental phenomena such as believing, intending, and wanting) has to be conceptually linked with public phenomena. Otherwise, to put the matter in the terminology of behavior analysis, the verbal community could not teach children to use such terms appropriately. The psychological terms must have some external referent in preverbal behavior. But, on the other hand,

> . . .the employment of psychological terms outstrips their foundation in preverbal behavior. Someone who has satisfied us that he understands certain psychological terms begins to use them in first-person statements *in the absence* of the primitive, preverbal behavior that had previously served as the basis for judging that he understood those terms. He tells us that he feels ill, or angry at someone, or worried about something when we should not have supposed so merely from his demeanor. The interesting point is that in a great many cases we will *accept* his testimony. We conclude that he is angry when, if we had been judging solely on the basis of nonverbal behavior and visible circumstances, we

should not have thought it. We begin to use his testimony as a new criterion of what he is feeling and thinking, over and above and even in conflict with the earlier nonverbal criteria. (Malcolm, 1977, p. 101)

Nor does radical behaviorism have an adequate answer to the need to find a third-personal level of analysis in order to make sense of observed behavior. This is essentially the need for a heterophenomenological methodology for reconstructing the beliefs, attitudes, desires, and motives that would render such observations intelligible.[3]

The Continuity of Behavior

The plausibility of an extensional radical behaviorist interpretation depends vitally upon its capacity to account for the continuity of behavior. Why should behavior that has been followed by a particular ("reinforcing") stimulus in the presence of a setting stimulus be re-enacted when a similar setting is encountered? Why should a rule that describes certain physical or social contingencies be followed at some future date when those contingencies are encountered? Why can I tell you now what I ate for lunch yesterday? The whole explanatory significance of learning history is concerned with the continuity of behavior between settings, and this implies some change in the organism—some means of recording the experience of previous behavior in such a way that it will be available next time similar settings are encountered. There is no other way in which the individual can recognize the potential offered by the current behavior setting in terms of the reinforcement and punishment signaled by the discriminative stimuli that compose it.

The radical behaviorist account of behavioral continuity requires that a common stimulus or some component thereof is present on each occasion that a response is emitted. The stimulus must be either a learned discriminative stimulus and/or a reinforcer. The difficulty with this is that it is not always possible to detect each element of the three-term contingency when behavior is learned or performed. The tendency is, then, to suppose that something occurs within the individual, presumably at a physiological level, that will one day be identified as sufficient to account for the continuity of behavior—but the problem is less one of ontology than of methodology, of the theoretical imperatives involved in explaining the continuity of behavior and therefore the language employed to account for it.

The issue revolves around *what is learned*. Whether one assumes that learning takes place as a result of initial exposure to a reinforcing stimulus and that behavioral control is transferred contingently to a paired setting stimulus that acquires discriminatory significance (the standard radical behaviorism view) or that learning usually occurs as a result of observing a conspecific's behavior and its consequences, the only way in which such learning can be described requires

---

[3] Although Dennett (1991) presumably intends adoption of a third-personal heterophenomenology to exclude a first-personal phenomenological position, there is no reason why the separate adoption of latter is precluded by acceptance of the former.

the use of intentional idioms. A purely descriptive account can, where this is possible, relate responses to the stimuli with which they correlate, and by which they are therefore predictable and open to influence. This is the essential program of an extensional behavioral science, and I do not wish in any way to argue that it be other than enthusiastically executed. Indeed, it is important to my research program that it is. However, I would like to discuss the possibility that it is not always feasible to make the required connections between environment and behavior, and that this acts as a stimulus to the discovery of an explanation rather than a mere description of behavior and its contextual determinants. The quest for explanation will always be there, should behaviorists choose to adopt it, but the failure of the extensional approach is a catalyst to its implementation. Let us examine these two separable claims in turn.

The Incompleteness of the Extensional Account

Although he does not use the term "behavioral continuity," Bandura (1986) provides a clear description of the problem. The arguments against radical behaviorism he puts forth center on the impossibility of providing an account of behavioral continuity that does not refer to cognitive processing. So long as people are assumed to act automatically in response to the environmental consequences of their past behavior or their thoughts are conceptualized as no more than intervening events themselves under environmental control, so long will any "internal link in the causal chain" be eschewed and agency assumed to reside in the environment (p. 12). Yet there are instances in which environmental causation is assumed to act without any apparent mechanism by which it produces behavior over time.

First, consider his treatment of the fundamental behaviorist principle that behavior is controlled by its immediate consequences. Bandura points to Baum's (1973) demonstration that the rate of emission of behavior is related to the aggregate of its consequences. Such "molar" behavior is actually a feature of Skinner's own approach since it is learning history rather than present stimuli alone that determine behavior. In fact, it was Herrnstein (1997) who most obviously defined and built upon this phenomenon. Defining choice not as an internal deliberative process but as a *rate* of intersubjectively observable events that are temporally distributed, Herrnstein's dependent variable was not the single response that needed contextual explication in terms of a single contingent reinforcer—it was the relative frequency of responding, which he explained by reference to the relative rate of reinforcement obtained from the behavior. Animals presented with two opportunities to respond (e.g., a pigeon pecking key A or key B), each of which delivers reinforcers (brief access to grain) on its own variable interval (VI) schedule, allocate their responses on A and B in proportion to the rates of reinforcement they obtain from A and B. This phenomenon, known as "matching," has been replicated in numerous species, including humans, and has found applications in behavior modification and organizational behavior management, to name but two relevant fields. In particular, it provides a framework for the

behavioral analysis of consumption (Rachlin, 1989, 2000). However, Baum's (1973) molar approach, to which Bandura makes reference, is sufficient to suggest that organisms are sensitive to how often a response is reinforced over a long period of time and that their behavior is thereby regulated according to the aggregate level of reinforcement. Such integration, Bandura asserts, requires cognitive skills and actually suggests the need for a subsumptive level of analysis—cognitive, environmental, physiological, behavioral, or otherwise—such as that called for by Smith (1994). The absence of any convincing evidence for these (when cognition is given a specific ontological status) leaves the ascription of intentional content as the only safe possibility given the current state of knowledge.

A second consideration to which Bandura draws attention is that when behavior is learned on intermittent schedules, only a small proportion of responses receive reinforcement, and reinforcements are occasional—perhaps only every 50th or 500th response is reinforced. Yet the behavior may strengthen over very long periods. Similarly, extinction may be accordingly prolonged. The question is whether such integration or behavioral continuity can be explained without positing some non-environmental determinant, presumably cognitive. Something other than external causation is necessary to account for what happens in between. Bandura invokes the distinction between the acquisition of a skill and its performance which, in turn, evokes the question of what is learned. Cognitive processes are again implicated. Despite (or because of) the fact that the delivery of reinforcement may be highly intermittent on such schedules, consistent patterns of behavior are acquired that define the continuity of behavior.

Bandura also points out that most complex behavior is learned by modeling rather than by experienced reinforcement (1986, pp. 74-80). He is highly critical of operant attempts at interpreting observational learning within the framework of the three-term contingency, which portray the process as one in which the modeled stimulus (S^D) is followed by an overt matching response (R) which produces a reinforcing stimulus (S^R). The elements of the three-term contingency are often missing from actual instances of observational learning. When the observer performs the matching response in a setting other than that in which it has been modeled—when neither the model's behavior nor that of the observer is reinforced, and when the modeled behavior is performed by the observer after the passage of time (which may be several months)—the operant paradigm is unable to explain the behavior. As Bandura (1986, p. 74) points out, "Under this set of conditions, which represents the pervasive form of observational learning, two of the factors ($R \rightarrow S^R$) in the three-element paradigm are absent during acquisition, and the third factor (S^D, the modeling cue) is absent from the situation in which the observationally-learned behavior was first performed." Observational learning of this kind also requires some mechanism to aid integration of vast amounts of information. Acquisition of novel behavior in particular requires such integration of modeled information. Bandura maintains that learning through modeling requires four processes: attentional, retentional, reproductive, and motivational. Certainly, observational learning is a process that must be comprehended at the personal level of analysis. Neither sub-personal nor super-personal levels can cope with it.

Learning that involves rule acquisition and following must also require these four procedures in some way or other. The individual acquiring rules from others must pay attention to the behavior of others, verbal or nonverbal. Somehow this has to be retained and compared, for instance, with earlier-gained knowledge and experience. Then it must somehow be translated into overt behavior when there is situational immediacy that makes the behavior in question possible or even likely. For Bandura, all of this argues for cognitive representation and processing, and it becomes all the more urgent to develop this line of reasoning if *understanding* rather than prediction and control is the primary goal of scientific endeavor. But whether the inclusion of cognitive processing will increase the predictability of behavior is an empirical question. In fact, we must keep an open mind on whether invocation of cognitive mechanisms adds to predictive accuracy. Their primary aim is to aid understanding, to allow a complete account of human behavior acquisition and maintenance. The environmental variables alone *might* contribute more to simple prediction and control; however, the evidence is that cognitive factors alone add little to prediction (Foxall, 1997).

Can the required account of behavioral continuity be achieved by introducing the moderating effect of thought into the explanatory scheme? Bandura argues that a fundamental principle of radical behaviorism is that thought cannot affect action. He argues that, contrary to this, most external influences on behavior act via cognitive processing. People develop beliefs about what is happening to them (i.e., the likely consequences of their behavior) and the beliefs come to influence their behavior. Moreover, "One can dispense with the so-called internal link in causal chains only if thought cannot affect action" (1986, p. 13). It is a moot point, however, whether thought influencing behavior is or is not part of radical behaviorist explanation. Strictly, thought is a collateral response, the effect of the same environmental events that determine the overt responses with which the thoughts are associated. However, even Skinner came to recognize thoughts and other private events as "non-initiating" causes in the sense that they might act as discriminative stimuli for covert and overt behaviors but remained ultimately dependent on external environmental stimuli for their power (as did the events of which they were local or proximal causes). Other radical behaviorists have held that a private event can function as any of the elements in the three-term contingency—hence, a thought can reinforce other covert or overt behaviors, though this remains a subject of deep controversy. More particularly, however, the role of thought in rule-governed behavior is of interest here. Rules may inhere in thought, and thought, like other verbal behavior that embodies or expresses rules, may thus control responses. This is an interesting departure from the behaviorist view that behavior can predict other behavior but not be the cause of it.

What Is Learned?

Dennett comments that "The difficulty the behaviorist has encountered is basically this: while it is clear that an experimenter can predict rate of learning, for example, from the initial conditions of his mazes and experience history of his

animals, how does he specify just *what* is learned?" (1969, pp. 33-34). "What [the animal in the maze] learns, of course, is *where the food is*, but how is this to be characterized non-Intentionally? There is no room for 'know' or 'believe' or 'hunt for' in the officially circumscribed language of behaviorism; so the behaviorist cannot say that the rat knows or believes that his food is at x, or that the rat is hunting for a route to x" (p. 34).

Considerations such as these have led behavioral scientists to theorize about the nature of learning. Mediational theories such as those of Hull and Tolman have given way to the use of intentionality to explain behavior, not on the basis of positing intervening variables but as an inevitable linguistic turn. Berridge (2000) makes the progression from mediationism to intentionalism clear in his description of the history of behavioral psychology. Bolles's (1972) account of behavior in terms of the expectation of hedonic consequences follows the S-S theory of Tolman rather than the S-R theory of Hull but suggests that what are learned are S-S associations of a particular kind and function: an association is learned between a conditioned stimulus (CS) and a subsequent hedonic stimulus (S*) that elicits pleasure. The first S does not elicit a response but an expectation of the second S (S*). Bolles (1972) developed a "psychological syllogism" in which, as Dickinson puts it:

> Exposure to stimulus-outcome (S-S*) and response-outcome (R-S*) contingencies leads to the acquisition of S-S* and R-S* expectancies, respectively, representing these associative relations. The two expectancies are "synthesized" or combined in a "psychological syllogism" so that in the presence of the cue, S, the animal is likely to perform response R. (1997, p. 346)

The response becomes more probable as the strengths of the expectancies increase and as the value of S*, which is influenced by the animal's motivational state, increases. Bolles employs this theory to explain why animals sometimes act as though they have received a reward when they have not: for example, the raccoon that washes a coin as though it were food, "misbehavior," autoshaping, or schedule-induced polydipsia—all empirical instances that research in the 1960s and 1970s had shown to be contraindicative of the reinforcement model.

Berridge (2000) argues that useful as this is, it fails to explain why the animal still approaches the reinforcer (say, food) rather than waiting for it to appear and enjoying the S* in the interim. He discusses the approach of Bindra (1978), who proposes the hedonic transfer of incentive properties to the CS. Bindra accepts the S-S* theory but argues that the S does not simply cause the animal to expect the S*—it also elicits a central motivational state that causes the animal to perceive the S as an S*. The S assumes the motivational properties that normally belong to the S*. These motivational properties are incentive properties that attract the animal and elicit goal-directed behavior and, possibly, consumption. Through association with the S*, the S acquires the same functions as the S*. An animal approaches the CS for a reward and finds the signal (S) attractive; if the CS is food, the animal wants to eat it. If it is an S for a tasty food S*, the animal may take pleasure in its attempt to eat the CS (Berridge, 2000, p. 236; see also Bouton & Franselow, 1997).

But if CSs were incentives, one would always respond to them whether or not one were hungry. The question is to explain how CSs interact with drive states. Toates (1986), therefore, builds on the Bolles–Bindra theory by positing that both cognitive expectancy and more basic reward processes might occur simultaneously in the individual. All of these theories are necessarily intentionalistic since they deal in expectancies.

Delimiting Behavioral Interpretations

The ubiquity of apparent three-term contingencies as we survey life beyond the lab raises difficulties for an interpretative account, which is meant to be more than "plausible." As radical behaviorism stands, its program of interpretative research is adjudged solely on the criterion of plausibility, thus there is no way of successfully delimiting the scope of its interpretations so as to meet the standards of validity and reliability decisive in qualitative as well as quantitative research.

This problem is inherent in Rachlin's (1994) extensional interpretation of observed behavior, *teleological behaviorism*, which proposes an interpretation of complex behavior based on final causes (i.e., the consequences of behavior). Final causes extend serially outward from the individual who behaves, each fitting or nesting into the pattern of the next. Hence:

> . . .eating an appetizer fits into eating a meal, which fits into a good diet, which fits into a healthy life, which in turn fits into a generally good life. The wider the category, the more embracing, the "more final" the cause. (Rachlin, 1994, p. 21)

The process of finding the causes of behavior is one of fitting the behavior into an ever-increasing molar pattern of behavior and consequences. Rachlin's system has no place for private events or intrapersonal phenomena; yet, unlike both radical and methodological behaviorism it freely employs mentalistic terminology. Rachlin asserts that mind is behavior—sequences or patterns of behavior rather than single acts. This molar view means that mental phenomena such as attitudes, intentions, and even pain are all defined by extended patterns of behavior. We know that our friend is in pain because of the behaviors he emits: grimacing, groaning, holding his arm, and so on—but this is not the central concern of this account. That lies in the fact that interpretations based on this system are unbounded and require an intentionalistic overlay of interpretation in order to be useful.

A whole series of final causes may each be nested within one another, diffused over time, with the whole sequence being necessary for a full explanation of the behavior that produced them. But since the events that explain a behavior are temporally extended, the compilation of its explanation may require the elapse of a significant period before the full complexity of the behavior's consequences can be noted and understood (Rachlin, 1994, pp. 31-32). The search for final causes as ultimate explanations may, nevertheless, be convoluted and unscientific in the sense that the propositions employed in explication of a behavior may never be brought into contact with the empirical events that could substantiate them or lead to their refutation. Rachlin's search for plausible extensions fails because the

extension identified is untestable (at least during the period of the interpretation). It is a travesty to say that the death of millions of Japanese civilians was a (or *the*) final cause of the physics research on atomic structure initiated by Rutherford and his colleagues. The two events are undoubtedly linked, but the invocation of a causal relationship between them is hardly adequate to account for either. More complete explanations must be sought at the intentional level.

Rachlin gives the example of our seeing a snippet of film showing a man swinging a hammer in order to explain Aristotle's conception of final causation (Rachlin, 2000, pp. 58-59; see also Rachlin, 1994, pp. 82-83, which is the subject of the review by Lacey quoted below). What, he asks, is the man actually doing? He might be swinging a hammer, hammering a nail, joining pieces of wood, laying a floor, building a house, providing shelter, supporting his family, being a good husband or father, or being a good person. All of these may be descriptions of his behavior, all may be true. But in order to arrive at the final judgment of what the man is doing we must look through the movie of the man's entire life: "The validity of any of the above descriptions may be settled by moving the camera back or showing more film—earlier and later" (2000, p. 59). The whole point of my criticism—the whole problem of the behavioral interpreter—is that there is no such thing as this comprehensive movie, no means of obtaining the complete behavioral history of this individual. We only get snippets of film and we need to find a means of interpreting it that is readier-to-hand than the supposed universal observation. Lacey hints at the kind of extra interpretation that is required:

> Insofar as building a house is constituted by an extended behavioral pattern, a particular act is part of the pattern only if it is performed because the builder believes that it will contribute towards her goal of a house being built through her own agency. In this analysis, which is Aristotelian, intentional categories are essential for defining the behavior pattern. The applicability of intentional categories to states of an organism cannot be grounded in the operant processes of discrimination alone. (1995/1996, p. 69)

The behavior of the builder is predictable only insofar as we ascribe to her the desire to build a house and the belief that placing this brick will lead to building a wall, that building the wall will contribute to the fabrication of a room, and so on. We need some mechanism for attributing these desires and beliefs. We do so partly on the basis of an idea of rational behavior in the circumstances and partly by inference from the builder's behavior pattern, including her verbal behavior. This displays an initial analysis based on the contextual stance (the operant behavior she displays is likely to result in this consequence that will increase the probability of her doing such and such next. . .), which is overlain by an intentional heuristic based on optimality, the assumption of desires and beliefs appropriate to the situation, and the ruling out of consequences that are improbable or non-maximizing (i.e., are impossible ends to attach to her behavior). We can immediately rule out the possibility that she is building a staircase to heaven, therefore, or a marble palace, or a headquarters building for the Society for the Protection of Small Disingenuous Wooden Italians with Extensible Noses. But can

teleological behaviorism? It may be significant that the predictive capacity of this approach is greatest in the context of the narrowly-defined and measurable utility functions of everyday economic life rather than in broader circumstances.

Statements couched in intentional terms are intentionally opaque: they are not susceptible to coextensional substitution (Juarrero, 1999, p. 38). The alternative consequences that could be spun out indefinitely by teleological behaviorism are coextensives in an intentional sentence: "I intend to place one more brick/build a wall/construct a cathedral." The intentional form emphasizes that these alternatives cannot be substituted one for another on the basis of the information the interpreter has as he makes his observation of what the person is doing.

Ascribing Intentionality

It is one thing to propose that a psychological theory requires the ascription of content to a pre-existing extensional theory such as that provided by physiology, but quite a trickier endeavor to justify the proposed level of analysis in psychologically-relevant terms, lay down procedures for the process of ascription, and specify the relationship between the two. So Dennett (1969) finally explores the manner in which content might be ascribed to physiological systems. This is the crux of his argument against behaviorism and for a psychology that proceeds by ascribing content to the extensional facts of physiology. The personal level of explanation (which is Dennett's focus here and which he contrasts with the sub-personal level at which physiology operates) is that of "people and their sensations and activities" rather than that of "brains and events in the nervous system" (p. 93). The sub-personal level provides mechanistic explanations, but they are not appropriate to the explanation of so-called mental entities such as pain. While there is a good understanding of the neurological basis of pain, Dennett raises the question of whether the presumed evolutionarily-appropriate afferent–efferent networks underlying this understanding are sufficient (they are certainly necessary) to account for the "phenomena of pain." This resolves itself into the question of whether pain is an entity that exists in addition to the physical questions that constitute this network (p. 91).

There are no events or processes in the brain that "exhibit the characteristics of the putative 'mental phenomena' of pain" that are apparent when we speak in everyday terms about pain or pains. Such verbalizations are non-mechanical, while brain events and processes are mechanical. It is unclear, for instance, how an individual distinguishes a sensation of pain from a nonpainful sensation. The only distinguishing feature of pain sensations is "painfulness," which is an unanalyzable quality that allows for only circular definition. But people can do this, and the personal level is the level at which pains are discriminated, not the sub-personal. Neurons and brains have no sensation of pains and do not discriminate them. Pains, like other mental phenomena, do not refer. Our speaking of them does not pick out *any thing*; pain is simply a personal-level phenomenon that has, nevertheless, some corresponding states, events, or processes at the sub-personal, physiological level. This is not an identity theory; Dennett does not identify the

experience of pain with some physical happening. He maintains two separate levels of explanation: one in which the experience of pain, while felt, does not refer, and one in which the descriptions of neural occurrences refer to actual neural structures, events, and states in which the extensionally-characterized science deals.

The task now becomes that of ascribing content to the internal states and events. The first stage is straightforward since intentional theory assumes that the structures and events they seek to explain are appropriate to their purpose. An important link in this ascription is provided by hypotheses drawn from the natural selection not only of species but of brains and the nervous system—a system which, through evolution, has the capacity to produce appropriate efferent responses to the afferent stimulation it encounters. It clearly has the ability to discriminate among the repertoire of efferent responses it might conceivably make. Its ability to discriminate and respond to the stimulus characteristics of its complex environment means that it must be "capable of interpreting its peripheral stimulation" to engender inner states or events that co-occur with the phenomena that arise in its perceptual field. In order for us to be justified in calling the process intelligent, something must be added to this afferent analysis: the capacity to associate the outcomes of the afferent analysis with structures in the efferent portion of the brain.

For instance, in order to detect the presence of a substance *as food*, an organism must have the capacity not only to detect the substance but, once encountered, to stop seeking and start eating. Without this capacity to associate afferent stimulation and efferent response, the organism could not be said to have detected the presence of the substance *as* that of food. Dennett uses this point to criticize behaviorists for having no answer to the question how the organism selects the appropriate response. There is a need to invest the animal that has discriminated a stimulus with the capacity to "know" what its appropriate response should be.[4]

The content of a neural state, event, or structure relies on its stimulation *and* the appropriate efferent effects to which it give rise, and in order to delineate these it is necessary to transcend the extensional description of stimulus and response. It is necessary to relate the content to the environmental conditions as perceived by the organism's sense organs so that it can be given reference to the real-world phenomena that produced the stimulation. It is equally important to specify what the organism "does with" the event or state so produced in order to determine what that event or state "means" to the organism. An aversive stimulus has not only to be identified along with the neural changes it engenders to signify that it means danger to the animal; in addition, the animal has to respond appropriately to the

---

[4] In fact, behaviorists have ducked this problem by designating it as part of the physiologist's assignment and drawing the conclusion that behavioral scientists need be concerned with it no longer. The conventional behaviorist wisdom over the kind of cognitive ascription to which Dennett refers is that it amounts to no more than "premature physiology."

stimulus (e.g., by moving away). Failure on its part to do so would mean that we were not justified in ascribing such content to the physiological processes occurring as a result of the stimulation. If we are to designate the animal's activities as "intelligent decision making," then this behavioral link must be apparent. Only events in the brain that appear appropriately linked in this way can be ascribed content and described in intentional idioms.

How, then, are the intentional ascription and the extensionally descriptions related? This ascribed content is not an additional characteristic of the event, state, or structure to which it is allocated; some intrinsic part of it discovered within it as its extensionally-characterized features are discovered by the physiologist. They are a matter of *additional interpretation*. The features of neural systems, extensionally characterized in terms of physiology or physics, are describable and predictable in those terms without intentional ascription, which makes reference to meaning or content. Such a scientific story, consisting in an account of behavior confined to talk of the structure and functions of neural cells and so on, is entirely extensional in character. But such an extensional story could not, according to Dennett, provide us with an understanding of *what the organism is doing*. Only an intentional account can accomplish this, "but it is not a story about features of the world *in addition to* features of the extensional story; it just describes what happens in a different way" (see Dennett, 1969, pp. 90-96). Such an extensional theory would be confined to the description/explanation of the *motions* of the organism rather than of its *actions*.

In practice, physiologists do not seem able to get along in their account of the function of the central nervous system without viewing neural operations as signals, reports, or messages (for modern corroboration see, for a typical textbook treatment: Gazzaniga, Ivry, & Mangun, 1998; and *inter alia* for direct research evidence: Angulo, Staiger, Rossier, & Audinat, 1999; Kandel, 2001). As Dennett puts it:

> Were the physiologist to ban all Intentional coloration from his account of brain functioning, his story *at best* would have the form: functional structure A has the function of stimulating functional structure B whenever it is stimulated by either C or D. . . .No amount of this sort of story will ever answer questions like why rat A knows which way to go for his food. If one does ascribe content to events, the system of ascription in no way interferes with whatever physical theory of function one has at the extensional level, and in this respect endowing events with content is like giving an interpretation to a formal mathematical calculus or axiom system, a move which does not affect its functions or implications but may improve intuitive understanding of the system. (1969, p. 79)

The required ascriptions of content would thus not comprise intervening variables within a physiological theory but a "heuristic overlay" on the extensional account. Such a centralist theory would have two components:

> . . .the extensional account of the interaction of functional structures, and an Intentional characterization of these structures, the events occurring within them, and states of the structures resulting from these. The implicit link between

each bit of Intentional interpretation and its extensional foundation is a hypothesis or series of hypotheses describing the evolutionary source of the fortuitously propitious arrangement in virtue of which the system's operation in this instance makes sense. (1969, p. 80)

The ascription of content to afferent and efferent operations is necessarily imprecise since it depends on the inexact locutions we use in everyday life.

Taking Stock

I have argued that the legitimate ascription of content relies emphatically upon the clear understanding of the nature of the personal level of analysis, a matter on which Dennett, after the brilliant insight of his 1969 essay, has proved somewhat flexible over the years. After the clear start he made in *Content and Consciousness* (1969), Dennett subsequently loosened the argument that content was ascribable only at the personal level of non-mechanical explanation and proposed that mechanistic, sub-personal systems may be treated as intentional systems in their own right via the ascription of content that permits their prediction. Dennett's introduction of the idea of sub-personal cognitive psychology is a particular source of confusion of his original personal/sub-personal dichotomy. The present paper now attempts, therefore, to lay an unambiguous basis for behavioral psychology, in which the level of analysis at which content ascription may properly be made is established as the personal level, and the range of non-personal levels of analysis which provide the basis for such attribution is extended from the sub-personal to the *super-personal* in order that behavior may be systematically incorporated in the scheme of explanation to which Dennett has given rise. It argues, moreover, that only at the personal level is it possible to ascribe intentionality in order to explain behavior. Of particular importance in the present context, it is the only level at which cognition and consciousness can be ascribed. The development in Dennett's thought with respect to the personal/sub-personal distinction is specifically addressed, and the confusion in philosophical psychology that has arisen from the proliferation of levels of explanation is examined. It is argued that the personal level, which is at least severely de-emphasized in Dennett's continuing work, is of central importance to the philosophical basis of cognitive psychology since it is the sole level at which cognition can be said to occur.

The Categorical Distinction

The personal level of explanation is of central importance to both intentional and behavioral psychologies—to the first because it is at this level alone that intentional content can be legitimately ascribed, and to the second because of the necessity of making sense of private events and their status in a causal theory. I do not think that anyone has argued more convincingly for the theoretical necessity of this level of analysis, nor done more to qualify its theoretical value, than Daniel Dennett. The grounds provided by Dennett for the legitimate ascription of content

have, with time, become vaguer insofar as they have become explicitly detached from the logic of evolutionarily-consistent reasoning that originally underpinned the distinction. Admittedly, the rules for ascribing beliefs and desires to a rational system in order to predict it (i.e., the execution of the intentional stance or, as Dennett refers to it, *the intentional strategy*) include consideration of what beliefs and desires such a system "ought" to have given its position and circumstances, and this is bound to include considerations that stem from its phylogenetic history as well as its current setting. But this is a less detailed and less logically constructed version of the procedure for allocation content than that given in *Content and Consciousness* (Dennett, 1969). This procedure involves the ascription of content to the theories and findings of the extensional sciences that proceed at the sub-personal level, but it leaves them intact in the process. There is an additional level of interpretation that does not take place in the terms of the extensional science on which it is built, thus it is not a contribution to neuroscience; rather, it comprises a heuristic device that is composed of intentional idioms, which do not belong in an extensional science. It exists on a level other than the sub-personal, which characterizes neuroscience. It exists on the personal level, the only level at which it is legitimate to ascribe content, according to the early Dennett. This is the process in which the ascription of intentional idioms takes place and the process that produces the personal level by prescribing in a way that is highly circumscribed (by the logic of evolution by natural selection) the content that an evolved entity "ought" to have by virtue of its phylogeny.

But there is more to the prediction and explanation of the organism's behavior than the ascription of content according to the principles of evolutionarily-consistent reasoning. The intentional strategy (Dennett, 1987) assumes that the behavior of an organism can be predicted only by the ascription of content relating not only to its evolution but also to its current position, those of its circumstances that signal the rewards and punishments of following a particular course of action *primed by* the organism's learning history gained in similar circumstances. It requires, in other words, the ascription of content (again to arrive at the personal level) on the basis of the theories and findings of extensional science, which deals with the effects of social and physical context on the ontogenetic development of the organism, including its acquisition of a behavioral repertoire. I submit that this science is behavior analysis, in which the fundamental unit of analysis is the environment–behavior contingency (Lee, 1988). Content may be legitimately attributed to the findings of this science on the basis of the principle of "selection by consequences" (Skinner, 1981), which includes not only natural selection but the process in which a behavioral repertoire is acquired in the course of operant conditioning. Behavior analysis thus provides an extensional basis for a *super-personal* level of analysis. How would the process of ascription be determined in this case?

Commentators on *Content and Consciousness* seem often to overlook the ambitious nature of Dennett's project: the resolution of the claims of extensional science with the inevitability of intentional explanation of behavior. This is to be achieved not by the super-conceptual integration of the two systems of

explanation, which retain their individual claims to uniqueness as explicators of one or other facet of natural phenomena, but to ground the intentional in a basis of materialism that rescues it from apparently unlimited phenomenological speculation.

Dennett notes the unfeasibility of S-R theorists—under which term he seems to subsume advocates of both respondent and operant behaviors—showing how a novel stimulus can arrive at or select the appropriate response. He points out, for instance, that an animal might detect a stimulus but not "know" what the appropriate response is (the stimulus in question could just as well be a discriminative stimulus as an unconditioned or conditioned one). No afferent can be taken by the brain to have a specific significance unless it is recognized by the efferent side of the brain has having just that significance (i.e., until the brain has produced the appropriate response). The content of a neural event or state depends not only upon its "normal state of stimulation" but also on whatever additional efferent effects it produces. The determination of these factors necessarily takes us beyond the extensional description of stimuli and responses. Indeed, as emphasized earlier, the lack of an account of behavior at the personal level imposes a severe restriction on radical behaviorist explanation (Foxall, 2004).

Content can accordingly be ascribed to a neural event when it is a link in an *appropriate chain* between afferent and efferent that has been selected in the course of the phylogeny of the organism in question. The content is not something to be discovered *within* this neural event; it is an extra interpretation, the rationale of which is not to understand better the operation of the subsystem *per se* but to provide a local justification for the ascription of appropriate content at the personal level. The ultimate justification for such ascription is provided by evolutionary thinking: the intelligent brain must be able to select the appropriate response to a specific stimulus. Why should this be less the case for the link between extensional operant analysis and the personal level of analysis than for the link between physiology and that level? Intentional ascription simply describes what a purely extensional theory would describe—nothing more—but in a different way. This different way may be useful to the physiologist, however. Neuroscience that does not view neural events as signals, reports, or messages can scarcely function at all. No purely biological logic can tell us why the rat knows which way to go for his food. Nor can any purely contextualistic logic reveal this in the absence of some sort of "Dennettian overlay." In neither case does the proposed intentional ascription detract from the extensional version of events; rather, it adds an interpretation that provides greater intuitive understanding of the system.

Hence, the sub-personal level is coterminous with that of an extensional science such as physiology, which is mechanistic in the explanations it provides. Intentional explanation simply *does not belong at this level* and we cannot add content to this level without violating its integrity as a conventionally scientific (i.e., extensional) approach to theory. We can, however, use it as the basis of appropriate content ascription (i.e., the attribution of intentional idioms that make certain behaviors of the organism intelligible—pain, for instance, or other emotional activity), but in so adding content we arrive at the personal level of

explanation. This is the sole level at which pain, etc. can be comprehended. There is a sharp epistemological dichotomy here between the personal and sub-personal levels of explanation, at one of which it is appropriate to include intentional explanations, the other serving as a basis for legitimately doing so but remaining intact as an extensional level of understanding. The guiding principle by which content is added is evolutionary logic. The process of natural selection that produced the findings identified at the level of physiology (or other sub-personal science) must provide the logic by which activities that are proposed in order to explain or predict the behavior of the whole organism.

The Intentional Stance

By the time of Dennett's (1981) distinction among three *kinds* of intentional psychology, however, his thinking demonstrates some subtle changes in the use of the terms personal and sub-personal levels and their relationships with psychology. He argues that folk psychology (the first kind of intentional psychology) provides a source of the other two: "intentional systems theory" and "sub-personal cognitive psychology." Folk psychology provides a non-specific and unhelpful causal theory of behavior; a more systematic and useful predictive tool requires refinement. The distinction between logical constructs or *abstracta* and causally-interacting *illata* provides a key. While the beliefs and desires of abstract and instrumental folk psychology are abstracta, the interactive theoretical constructs of sub-personal cognitive psychology are illata.

Each of the two additional intentional psychologies Dennett proposes rests integrally on one or other. *Intentional systems theory* (the second kind of intentional psychology) draws upon the notions of belief and desire but provides them with a more technical meaning than they receive in folk psychology. It is a whole-person psychology, dealing with ". . .the prediction and explanation from belief–desire profiles of the actions of whole systems. . . .The *subject* of all the intentional attributions is the whole system (the person, the animal, or even the corporation or nation [see Dennett, 1976] rather than any of its parts. . .)" (Dennett, 1987, p. 58). Intentional systems theory is a competence theory in that it specifies the functional requirements of the system without going on to speculate as to what form they might take (Dennett, 1978, p. 74). The necessity of this general-level theory is that of providing an account of intelligence, meaning, reference, or representation. Intentional systems theory is blind to the internal structure of the system. According to Dennett, the capacity of abstracta to interrelate, predict, and partly explain behavior itself suggests some underlying mechanism to which intentional systems theory does not, in principle, address itself. Any intentional system of interest would surely have a complex internal structure, and chances are that this will be found to resemble closely the instrumental intentional interpretation. The third kind of intentional psychology, "*sub-personal cognitive psychology*," is tasked with explaining the brain as a *syntactic* engine as opposed to the task of intentional systems theory, which is to explain it as a *semantic* engine (see below). By contrast to intentional systems theory, which is a competence

model, sub-personal cognitive psychology is a performance model, one designed to specify the workings of the system in terms of variables that can enter into a scientific theory, that is, in terms of illata (Dennett, 1978, pp. 74-78).

Note that this terminology and leap in analysis is itself misleading and confusing. The level of illata is simply that of cognitive psychology: there is no need to designate it further as *sub-personal*. This immediately suggests a source of confusion with the sub-personal level of analysis, especially as it has been defined in terms of the extensional rather than the intentional. It appears, moreover, that Dennett is thinking of cognitive psychology as requiring no other qualification than sub-personal. He does not contrast this with super-personal cognitive psychology, for instance—this does not enter into his thinking at all. The distinction he is making is between the personal level, at which abstracta operate, and that of another level at which illata operate. For him, this is the sub-personal. But, as I aim to show in this paper, cognitive psychology may require sub-personal *and/or* super-personal qualification, depending upon which source of justification for the ascription of cognitive variables at the personal level of analysis is being alluded to. In the case of sub-personal cognitive psychology this remains the naturally selected afferent–efferent linkages embedded in neurology; in the case of super-personal cognitive psychology it is environment–behavior linkages. I shall argue, further, that both abstracta and illata are phenomena of the personal level. The analysis appropriate to what is internal to the organism is neuroscience; intentionalistic psychological theories must perforce deal at the personal level.

The underlying mechanism to which even abstracta appeal in their capacity to explain and which (sub-personal) cognitive psychology attempts to uncover and explicate must surely be physiological in nature and requires an extensional neuroscience to reveal it. Cognitive psychology remains as theoretical an enterprise as intentional systems theory: its focus is the explanation of the brain as syntactical through its identification of the cognitive variables that may be legitimately ascribed at the personal level. In doing this it draws upon both its sub-personal and super-personal inputs—neurology and behavior—both of which are approached through an extensional science. But cognitive psychology itself remains an intentional science. To deviate from this "pure" dichotomy of the personal and sub-personal/super-personal is to invite the "mereological fallacy," the attribution to parts of a system of features that properly belong only to the system as a whole (Bennett & Hacker, 2003; see box on the next page).

Mereological Confusion

The presupposition Dennett makes about sub-personal cognitive psychology (to the effect that elements of sub-personal physiology such as the brain and the remainder of the nervous system can be treated as intentional systems in themselves in order to predict their behavior) certainly comes under considerable criticism from Bennett and Hacker (2003) on the grounds that it is intelligible to ascribe intentionality only to systems that are known to bear beliefs and desires.

Bennett and Hacker (2003, p. 73) specifically criticize Dennett on the grounds that he has committed the "mereological fallacy." Mereology refers to the logical relations of parts and wholes. The mereological principle says that psychological predicates that properly apply only to human beings cannot be meaningfully attributed to parts of the human being such as the brain. Hence, the ascription to a part of an organism the attributes that properly belong only to the whole organism is the mereological fallacy: "Human beings, but not their brains, can be said to be thoughtful or thoughtless; animals, but not their brains, let alone the hemispheres of their brains, can be said to see, hear, smell and taste things; people, but not their brains, can be said to make decisions or to be indecisive." (ibid). This fallacy is widespread among neurophysiologists, they claim, citing scientists who argue that "the brain *has experiences*, *believes things*, *interprets* clues *on the basis of information* made available to it, and *makes guesses. . .categorizes. . .*and *conceptually manipulates rules. . . .*[T]he brain *knows* things, *reasons* inductively, and *constructs hypotheses* on the basis of arguments, and its constituent neurons are *intelligent*, can *estimate probabilities*, and *present arguments*. [T]he brain *poses questions*, *searches for answers*, and *constructs hypotheses. . . .*[B]rains *decide*, or at least 'decide', and *initiate voluntary action.* [T]here are *symbols in the brain*, and the brain *uses*, and presumably *understands*, *symbols. . . .*[T]he brain *makes classifications*, *comparisons*, and *decisions*" (pp. 68-70). However, these authors contend that we simply do not know what it is for brains to do these things: we only know what it is for a person to do them. They further point out that whether psychological behavior can be imputed to the brain is a philosophical matter rather than an empirical one.

The mereological principle is non-empirical: it is not subject to confirmation or disconfirmation by experimentation or other empirical means. It is a convention and could be overturned, but only at the cost of changing a great deal else—changing the meaning of words and a host of familiar concept (p. 81). These authors are, therefore, adamant that intentionality cannot be ascribed to artifacts like computers, thermostats, molecules, brains: "Not only is it a subclass of psychological *attributes* that are the appropriate bearers of intentionality and not animals or things, but, further, only animals, and fairly sophisticated animals at that, and not parts of animals, let alone molecules, thermostats or computers, are the subject of such attributes" (p. 423).

This seems equivalent to saying that there has to be an ontological basis for a being/entity's believing in order to ascribe rational beliefs to it. The problem with Dennett's levels arises when he switches from the sub-personal as a means of working out what intentional idioms can be ascribed at the personal level to the idea that the *raison d'être* of intentional idioms is their capacity to predict behavior. He then gets involved in how to predict sub-personal events, characterizes them as (parts of) intentional systems, and applies the intentional stance.

"Given Dennett's conception of the intentional stance, it is unclear what precisely he means by claiming that the brain gathers information, anticipates things, interprets the information it receives, arrives at conclusions, etc. Presumably *he* is 'adopting the intentional stance' towards the brain, and is treating it as if it were a rational agent that believes what it ought to believe and desires what it ought to desire and acts on its beliefs and desires. But this is not coherent. We know what it is to treat a young child as if it were an adult, rational human being, but so we have any idea what it would be like to treat a *brain* as if it were a rational being? The brain. . .is not a possible subject of beliefs and desires; there is no such thing as a brain acting on beliefs and desires, and there is nothing that the brain does that can be predicted on the basis of its beliefs and desires" (Bennett & Hacker, 2003, p. 426). This is precisely what Dennett does in his later work: treating the subsystem of the brain as an intentional system. The answer is to restrict (a) the ascription of intentionality to the personal level *of people*, (b) to use the intentional stance again only for persons, and (c) to use the intentional stance as a device for explanation rather than prediction (cf. Elton, 2003).

The subtle difference that accompanies Dennett's later usages inheres in his argument (1978, p. 154) that the behavior of the person as a whole is the outcome of the interactive behavior of its various subsystems (Hornsby, 2000, pp. 16-17). This is a departure from his earlier insistence that to move to the sub-personal level (i.e., to the operation of the central and peripheral nervous systems) is to leave behind the personal level of explanation of sensations, intentionality, and behavior. Hornsby argues that this is inconsistent with the proscription of using sub-personal level findings to understand the personal level. Why-questions about the behavior of an actor in an environment can be answered only at the personal level. The problem is Dennett's later claim that the program of sub-personal cognitive psychology is to show how the physicalist findings of sub-personal extensional science can be used to interpret a fully-realized intentional system operating at the personal level.

Hornsby (2000, pp. 20-21) seeks to maintain the distinction between personal and sub-personal levels of explanation by arguing that intentional phenomena are real at the levels of persons but merely as-if constructions at the sub-personal level. The attraction of this distinction is that it maintains the independence of the personal level as a basis of explanation but permits the intentional stance to be operated at the sub-personal level for purposes of predicting the behavior of subsystems. Her somewhat arbitrary assertion, however, makes an ontological distinction between persons and non-persons (animals, subsystems, machines) that Dennett has sidestepped by adopting an instrumental approach to the ascription of intentionality (ascription is justified if it facilitates prediction; to have a mind or its constituent beliefs, desires, etc. is simply to have behavior that is predictable on the assumption that one has a mind, beliefs, desires, etc.). Dennett seems to have dealt with her position already (1978, pp. 272-273). Again, it is the distinction between Dennett's earlier and later positions on instrumentalism versus realism that seems to be at the heart of the misunderstanding.

This problem clearly rests on the more limited project in *Content and Consciousness*, which focused on the materialistic delineation of the intentional by reference to a genuinely sub-personal level of neuroscience, and Dennett's later enterprise focusing on the development of sub-personal cognitive psychology as the basis of a philosophy of cognitivism. In *Content and Consciousness* Dennett was concerned primarily with establishing how the findings of an extensional neuroscience could be brought into a single framework of analysis along with the conclusions of a mental science which recognized, albeit critically, the reality of intentional phenomena. The overall framework was to be consistent with materialism and thus to eschew an easy metaphysical reconciliation. The aim was to devise an a-ontological basis for the systematic ascription of content to the findings of physiology (a science of the sub-personal), the purpose of which was not to enhance or contribute to neuroscience *per se* but to justify the intentional ascriptions that could be legitimately made to the whole person. What could this physiology be said to have by way of beliefs and desires? The ascriptions are not part of physiology but an additional interpretation that makes possible analysis at

the personal level that prevents the unwarranted ascription of intentional content simply to explain behavior on mentalistic, uncritical, folk-psychological grounds.

The implications for cognitive psychology that are drawn in a rudimentary way in *Content and Consciousness* are that it is an activity of the personal level that inextricably involves intentional phenomena, but that it has been reconciled with extensional science through the agency of natural selection. It is when Dennett (1978) seeks to clarify the nature of intentional psychology by introducing a distinction between intentional systems theory and sub-personal cognitive psychology that potential confusion arises. The difficulty revolves around the status of sub-personal cognitive psychology.

In seeking to resolve it, it is important to appreciate that the sub-personal level of explanation that Dennett (1969) describes differs from the sub-personal cognitive psychology he introduced later (1981). While the sub-personal level consists of extensional science, sub-personal cognitive psychology, being derived from folk psychology, adds to this the heuristic overlay of ascribed intentional content that properly belongs at the personal level. However, many levels of explanation emerge from Dennett's scheme, the most prevalent current cognitive psychology, *social cognition*, is founded upon another: the super-personal level, an extensional psychology founded upon the "contextual stance," which relates behavior to the environmental contingencies that shape and maintain it (Foxall, 1999) and which is linked to the personal level of explanation by means of the ascription to the theories and findings of that extensional science of intentional content (Foxall, 2004). The three stances (pre-empirical methodological prescriptions for the prediction and partial explanation of systems) that Dennett proposes provide a key to the number of levels of explanation at which psychology may operate and clarify some of their interrelationships. In particular, the design stance, which Elton (2003) shows to be bifurcated in Dennett's writings, is highly relevant to the distinction between the sub-personal level of explanation and sub-personal cognitive psychology.

Sub-Personal Cognitive Psychology

As useful as a competence theory is, there has to be some underlying internal structure that accounts for the capacity of the various abstracta that are the components of intentional systems theory to predict systemic behavior at the personal level so well. Discovering this structure and its workings is the task of the third kind of intentional psychology: sub-personal cognitive psychology, the task of which consists in "[d]iscovering the constraints on design and implementation variation, and demonstrating how particular species and individuals in fact succeed in realizing intentional systems" (Dennett, 1987, p. 60).

The task of the brain, according to intentional systems theory and evolutionary biology, is semantic: it must decipher what its stimulus inputs mean then respond with appropriate behavior. But, in fact, to the physiologist the brain is no more than a syntactic engine: it "discriminate[s] its inputs by their structural,

temporal, and physical features and let[s] its entirely mechanical activities be governed by these 'syntactic' features of its inputs" (Dennett, 1987, p. 61). Hence:

> . . .it is the task of sub-personal cognitive psychology to propose and test models. . .of pattern recognition or stimulus generalization, concept learning, expectation, learning, goal-directed behavior, problem-solving—that not only produce a simulacrum of genuine content-sensitivity, but that do this in ways demonstrably like the way people's brains do it, exhibiting the same powers and the same vulnerabilities to deception, overload and confusion. It is here that we will find our good theoretical entities, our useful *illata*, and while some of them may well resemble the familiar entities of folk psychology—beliefs, desires, judgments, decisions—many will certainly not. . . .The only similarity we can be sure of discovering in the *illata* of sub-personal cognitive psychology is the intentionality of their labels (see *Brainstorms* [Dennett, 1978], pp. 23-38). They will be characterized as events with content, bearing information, signaling this and ordering that. (Dennett, 1987, p. 63)

> In order to give the *illata* these labels, in order to maintain any intentional interpretation of their operation at all, the theorist must always keep glancing outside the system, to see what normally produces the configuration he is describing, what effects the system's responses normally have on the environment, and what benefit normally accrues to the whole system from this activity. . . .The alternative of ignoring the external world and its relations to the internal machinery. . .is not really psychology at all, but just at best abstract neurophysiology—pure internal syntax with no hope of a semantic interpretation. Psychology "reduced" to neurophysiology in this fashion would not be psychology, for it would not be able to provide an explanation of the regularities it is psychology's particular job to explain: the reliability with which "intelligent" organisms can cope with their environments and thus prolong their lives. Psychology can, and should, work toward an account of the physiological foundations of psychological processes, not by eliminating psychological or intentional characterizations of those processes, but by exhibiting how the brain implements the intentionally characterized performance specifications of sub-personal theories. (Dennett, 1987, p. 64)

Evaluation

We should now appraise Dennett's altering conception and attribution of importance to the distinction of personal and sub-personal levels of explanation. Four distinct phases are apparent in his thought. The first is the so-called "categorical" distinction (held in varying forms by Davidson, 1980; Davies, 2000; Elton, 2000; Gardner, 2000; Hornsby, 2000), which maintains the analytical difference between these levels of explanation that Dennett set out in 1969. Dennett here holds to a strict personal/sub-personal distinction, using the latter to ascribe intentionality at the personal level. He also maintains a strict difference between extensional and intentional sciences, claiming that both are necessary. The role of behavior appears important here because it is to its explanation that the

ascription of intentionality is ostensibly directed. It receives no explicit definition or analysis, however; it is taken as a given, albeit an important one.

In the 1970s, and certainly by 1980, Dennett's criterion for the ascription of content changed from one that was explicitly justifiable on biological grounds to that of the predictability of behavior. This progression, by means of the introduction of the intentional stance, marks the abandonment of the personal level as a seriously entertained analytic category. The distinction between personal and sub-personal, crucial to the originally argued basis for the legitimate ascription of content, is lost as the intentional stance comes to be applied to sub-personal units in order to predict them (see also Hornsby, 2000). The mereological fallacy, inherent in Dennett's reasoning, rules out such a move despite the stand on realism that Dennett takes. Behavior is still important because its predictability is a criterion of the legitimate ascription of the mental—but it still receives no additional analysis, nor even a definition.

The third phase comes with Dennett's attempt to include cognitive functioning at the sub-personal level, the so-called "sub-personal cognitive psychology" that he has made the center of his philosophy of psychology. The sub-personal that is now the focus of attention is an intentional level of analysis that spans the divide between neurology and the personal. The categorical distinction is being further eroded. Behavior now is more sidelined than before. But is the notion of sub-personal cognition sustainable, or does cognition belong at the personal level?

The final phase (so far) is Dennett's explanation of consciousness. By now any suggestion that the personal is important appears to have been lost—though Elton (2000) disagrees—as the quest is for the heterophenomenological interpretation of behavior at the third-personal level. But Elton claims that consciousness can only be entertained at the personal level. Behavior again assumes importance if only because it must form the referential basis of heterophenomenological attribution of the content of consciousness. Without a firm basis for such attribution, however (in the form of rules for the legitimate ascription of intentional content that carry the same epistemological force that extensional neuroscience provided for intentionality advanced in *Content and Consciousness*), the use of mental idioms relapses into the loose mode of content attribution that Dennett's scheme in that work was specifically designed to guard against. I shall argue that the solution lies in the adoption of an extensional behavioral science paralleling that of extensional neuroscience. However, before exploring further the place of behavioral science in a comprehensive model of human behavior, two points about the sub-personal level of explanation remain to be made. The first arises out of Dennett's use of two apparently related terms, the "sub-personal level of explanation" and "sub-personal cognitive psychology." The latter term concerns the appropriate level of explanation to which cognitive psychology belongs.

The Sub-Personal Level and Sub-Personal Cognitive Psychology

The definition of the sub-personal level as being a target for intentional ascription (i.e., the view that the intentional stance can be used at any level of analysis at which it permits prediction superior to that provided by other stances), something which jars with Dennett's views in *Content and Consciousness* (though not with his later opinion), reflects the fact that Dennett uses the term "sub-personal" in two ways (Elton, 2000). In the first, the sub-personal level is devoid of ascribed intentionality in its own right. In *Content and Consciousness* he speaks of neuroscience in this capacity; the intentionality is something additional to the extensional science, and bringing them together results in the personal-level analysis by making legitimate (in terms of evolutionary reasoning) the ascription of appropriate content at that level. The extensional science is coterminous with the sub-personal level of analysis. In the second usage, the sub-personal level *incorporates* intentionality *per se*: this is the sub-personal cognitive psychology he speaks of in "Three Kinds of Intentional Psychology" (Dennett, 1981). There is no contradiction here since the sub-personal level he defines in *Content and Consciousness* and the sub-personal cognitive psychology he speaks of later are different levels of explanation, each of which draws upon the design stance in its own way (Elton, 2003, pp. 38-41). The first is the "causal blueprint perspective," in which the design of a system is closely related to its causal structure: whether the system is performing as designed can be ascertained by comparing its functioning with the blueprint of its causal structure. For instance, how well a chess computer is operating can be judged by comparing its operation with that specified by the program that regulates it. This design gives no indication of what the system is for; even a non-chess playing person could make a judgment of the efficacy of its operations in this manner. The design purpose of the machine does not attract consideration. The "teleological interpretation" of the design stance is, however, vitally concerned with the purpose of the system. The purpose that the system was designed to fulfill can be specified and its progress toward fulfilling that purpose monitored without reference to the causal blueprint of the system. We may know what a spark plug is designed to do and determine its success without knowing how it does it at the level of its sub-operations. We can nevertheless make predictions about the behavior of the spark plug and assess its efficacy in reaching its goal. In the case of the teleological perspective we are concerned with what the system as a whole is designed to do; with the causal blueprint perspective we are concerned with what the subsystems are designed to accomplish and with their interactions.

Acknowledging the two kinds of sub-personal analysis results, however, in the proliferation of levels of analysis and raises the question of where any particular psychological theory fits among them. This is quite clear in the case of social cognitive psychology, which is revealed to have connections with both the personal level and sub-personal cognitive psychology but to derive its distinctive purview of human behavior predominantly from yet another level of analysis, the super-personal level of explanation. The sub-personal level that Dennett identified

in *Content and Consciousness* is guided, in practice, by that design stance that Elton (2000) understands in terms of the causal blueprint perspective in which physiology attempts, by reverse engineering, to establish how the organism's subsystems have been designed in the process of evolution by natural selection. The physical stance is also apparent here. By contrast, sub-personal cognitive psychology involves the teleological perspective of the design stance in which intentionality is invoked and ascribed in the attempt to reverse engineer the system as a psychological entity in order to permit the prediction of the system as a whole.

Now Dennett claims that sub-personal systems can be treated as intentional systems (i.e., the intentional stance can be adopted toward them). Elton (2000, p. 4) notes that contrary to Dennett's initial clear distinction between the personal and sub-personal levels, he later (notably in *Consciousness Explained*, 1991) spoke as though the autonomy of the personal level were in doubt and as though an analysis in personal terms could be given of the sub-personal. He notes McDowell's (1994) claims that instead of maintaining the distinction between on the one hand (1) the relationship between a person and her environment and on the other (2) the relationship between different components of a person, Dennett conflates them. Hence, Dennett's claim that consciousness inheres in a person's capacity to access the content carried by a subsystem mixes stories that, must by their nature, be kept separate. According to Elton, however, Dennett does indeed want to pursue the idea of sub-personal processing as a prelude to his conclusions about consciousness, and the problem is reconciling this with the force of Dennett's original distinction between the personal and sub-personal. He claims that this is possible if pursued with understanding of how Dennett conceptualizes the ascription of content to systems and their component parts. Dennett has never denied that subsystems can be content-bearing, but he has not provided many clues as to how such ascription is to take place. Elton suggests that the procedure is as follows: Intentional states (attitudes towards contents) can be attributed, Dennett says, to intentional systems, which include people, frogs, chess machines, and robots. They can (Elton says of Dennett) also include *parts* of such systems, which can be seen as constituting smaller systems in themselves. On the assumption that the system under investigation has some goals (e.g., to survive, to win, to avoid injuring humans) and some "rational" means to achieve them (e.g., perceptual apparatus, powers of action), "*one then ascribes a whole network of intentional states (both cognitive and motivational) that best make sense of the system's behavior*" (Elton, 2000, p. 6). The constraint on this ascription process is knowledge of what intentional states the system "ought" to have, given its circumstances, and what, in view of the imperfections of the system, it is likely to have in actuality. This is the adoption of the intentional stance.

Elton proposes, as a means of comprehending so wide a range of systems and of remaining true to Dennett, that a content-bearing state be understood liberally as

> . . .no more than a state that is semantically evaluable and behavior guiding . . .what the state represents may or may not be the case and the behavior of the system will, in appropriate circumstances, be suitably affected by the presence of the state. Thermostats represent the temperature of the room—but of course a

thermostat may do so incorrectly if placed it in a draught—and this representation affects what they do, e.g. switch the boiler on or off. (Elton, 2000, p. 6)

So general a view of content requires that different kinds of content must be distinguished. Elton distinguishes content ascribed to a dog from that (less structured content) ascribed to a frog or (less structured still) that ascribed to a chess machine:

The content ascribed to persons has a structure such that it can figure in chains of reasoning, in expressions of justification and explanation, and so forth, the content ascribed to less cognitively sophisticated systems does not have a structure that is amenable to such uses. It does not, because there is no behavior that such systems can engage in that could count as, say, deliberative reasoning, justifying, or explaining. And ascription of content, in the view in question, cannot be divorced from the cognitive capacities of the system in question. (2000, pp. 7-8)

Elton also states that to adopt the intentional stance is to use "an autonomous level of explanation of the activity of that system" (2000, p. 8). The behavior of the person (say) whose behavior is predicted by the intentional stance is not being explained in terms of the component parts of that person qua intentional system, whether this is treated as an intentional system in itself or otherwise. Confusion arises because on the one hand, in the earlier Dennett, intentional ascription is something that can be justified only in terms of appropriate afferent–efferent linkages, while on the other, the later Dennett argues that intentional idioms, consonant with a broader understanding of the system's evolution and its present position, can be ascribed as long as they enable the system to be predicted. Although there may be an argument that any system that is so evolved and so placed that it can be predicted must have evolved the apt afferent–efferent linkages, this is too easy a way of overcoming the fact that we are presented in the earlier and later Dennett with essentially alternative devices for legitimately ascribing intentionality.

The complication that arises from this analytical uncertainty might be overcome in two ways: (1) by accepting that Dennett is speaking of sub-personal cognitive psychology when he says subsystems can be treated as intentional systems, or (2) by arguing that any system can be addressed with any of the stances he proposes and can, therefore, be studied as either a personal-level system or a sub-personal system. However, the conclusion is that whatever the merits of adding content *at* the sub-personal level, this action removes the justification in evolutionary terms of adding content to extensional science in order to arrive at what may be legitimately explained in intentional terms at the personal level. It opens the way for the identification of neurological activities as behavioral substrates, a rather different emphasis.

Cognition as a Personal Level Phenomenon

It is important to maintain the original personal/sub-personal distinction because cognition, whatever else it might be, appears to be a phenomenon of the personal level. This location is, moreover, crucial to our ideas of what cognition is, what determines it, and whether it is behaviorally causal. I argue that (1) the definition of cognition is such that it cannot be other than a personal level occurrence, and (2) what is normally understood as cognition overlaps sufficiently with Elton's understanding of consciousness, which he argues is a personal-level phenomenon, to make cognition locatable only at that level.

First consider what cognition is. Among many definitions from which we might select, Heyes (2000, p. 20) portrays it in terms of "theoretical entities providing a functional characterization of the operations of the central nervous system, which may or may not be objects of conscious awareness, and that are distinct from perceptual and motor processes." Theoretical entities are ascribed entities, and ascription belongs at the personal level. Cognitive terms are intentional and belong at the personal level. *Sub-personal* cognitive psychology attempts to link the intentional and the neurological—best if these are kept distinct conceptually for the reasons already given. The theoretical entities of cognition refer to functions of the central and peripheral nervous systems which, of course, make them material, but because they are entities of the kind exemplified by parallelograms of forces they are not physical in the sense that tables and neurons are. They are both invented entities and intentional objects; hence, they belong at the personal rather than sub-personal level. They are not part of an extensional science since they cannot be tested/verified in the same way as physically measurable entities. Therefore, they do not belong at the sub-personal level.

Second, consider Elton's argument that a distinction is in order between the intentional stance and the personal level/stance, both of which belong to the overarching category of rationalizing stances. The personal level arises when an individual can give reasons for his or her actions and when such reasons can be adjudged good or bad reasons. We can ask a consumer "Why did you buy the more expensive brand?" and receive the reply that "It will last longer" or "My friends expect me to." It does not follow that the behavior was actually motivated or caused by these reasons, only that they can be given in explanation. Some systems cannot give such reasons—animals and computers, for instance. We can use the intentional stance to predict their behaviors, but the personal stance can be taken only by humans. Elton's point is that consciousness is a property only of systems capable of providing the narrative accounts of their behavior. Such accounts require the ability to take the personal stance. Only such can reason, decide, deduce, and so on—or at least only such can describe their doing so. This can, I think, be tied into the personal phenomenology of thinking and knowing, feeling and emoting, that is part and parcel of what is generally called personal experience or consciousness. Unlike Dennett (1978), and perhaps like Skinner (1974), I do not deny its existence.

Taking Stock Again

Although Dennett refers often to behavior as that which is to be explained, he does not incorporate an extensional behavioral science into his scheme of explanation to counterbalance the extensional neurology on which he (at least originally) relies. Yet behavior is the criterion of the intentional and needs to be systematically related to its causal environment so that it may play its vital role in the framework of analysis proposed here. Radical behaviorism, which we have already explored as the extensional behavioral science *par excellence*, is now integrated to the development of intentional behaviorism. In order to fix the a-ontological nature of this philosophy of psychology, it is contrasted with a "super-personal cognitive psychology" that is in some ways its logical extension but not a necessary consequence of seeking a level of explanatory detail in the realm of intentionality. The remainder of this paper, therefore, proposes an alternative source of content ascription to that of Dennett's sub-personal neurology: super-personal environment–behavior linkages. In doing so it outlines an alternative strategy to the more usual philosophy-to-psychology approach that has often characterized the philosophy of psychology despite the naturalistic protestations of many of its exponents. If, however, instead of constructing our idea of the nature of social cognitive psychology on the basis of *a priori* philosophical assumptions we observe how social psychologists of cognition–behavior relationships operate, we shall the more easily find an adequate philosophical foundation for the scientific analysis of social cognition. The argument is that the ascription of content requires the incorporation of an extensional behavioral science as well as an extensional neuroscience; the position arrived at is a-ontological intentional behaviorism.

The Place of Extensional Behavioral Science

The intentional strategy provides only outline information on the attributions that are to be made in order to predict an intentional system. To say that these attributions consist in the beliefs and desires it "ought" to have is, of itself, little more than an abstract formula. Yet, the heterophenomenological device that Dennett (1982, 1991) proposes—in which the intentionality that "explains" an individual's behavior is reconstructed in much the same manner as that by which a literary critic analyses characterizations in a novel or play—invites a careful scrutiny of the system's prior behavior, applying the intentional stance to people's verbal behavior, treating it as a text to be interpreted in terms of their beliefs and desires. Much as one examines the text of a character in a novel in terms of what he or she says, what they do, and what others say of them plus background information about the author and his or her other writings, so one can produce an intersubjective account of the text provided by another person. The heterophenomenology of the person consists in an account of "*what it is like to be* that subject—in the subject's own terms, given the best interpretation we can

muster" (Dennett, 1991, p. 98). The resulting account is, like a scientific hypothesis, subject to testing in the face of the evidence, hence corrigible.

Now, this is only a highly generalized method of interpretation. It does not amount to a systematic means of incorporating behavior and its determinants into the framework. In order to accomplish that we shall need a more complete understanding of the requirements for a behavioral science, with its attendant means of interpretation and its role in a comprehensive framework for psychological research that incorporates both the intentional and the behavioral. For now, however, we face the immediate problem of determining what the link between behavior and ascribed intentionality might be in a more specific context.

Corrigible Behavior

A clue is that the correct beliefs and desires arise out of consideration of the logical consequences of the beliefs the system has previously held, which introduces issues that the contextual stance would attribute to learning history plus the current behavior setting and the consequences it portrays as contingent on behaving. The big problem is that Dennett does not give a concise and useable answer to the question of just which of the many available potential consequences of behavior are to be taken into consideration in this process. Dennett raises the problem that the logical consequences of behavior are infinite in number. Webb (1994) seeks the answer in the witnessed behavior of the system, a thesis he expands by discussing the three sources of attributable desires identified by Dennett (1987, p. 20) in what he described as his "flagship" exposition of the intentional stance.

First, there are the fundamental desires for basic goods: survival, the absence of pain, food, comfort, and entertainment—what some behaviorists might refer to as primary reinforcers. Second are those desires that can be rationally deduced from the system's goals and situation—perhaps conditioned reinforcers. Finally, Dennett introduces a form of witnessed behavior as a source of intentional ascription, one that is confined to consideration of verbal behavior. This raises problems of its own, however, since "The capacity to express desires in language opens the floodgates of desire attribution" (Dennett, 1987, p. 20). Dennett says little more about this means of belief attribution other than that these are beliefs the agent could not have had without language. As Webb (1994) points out, it would be useful here had Dennett made clear "just how we recognize utterances that call for desire attributions and how we know (based on what we hear) which desires to attribute" (p. 460). He notes, however, that Dennett's proposed method of intentional attribution allows for the attribution of desires on the basis of linguistic behavior. Dennett, moreover, is silent concerning any other attributions on the basis of witnessed behavior, but without some means of making these delimiting ascriptions of content any attributions made in order to operationalize the intentional stance are likely to be implausible.

If the intentional stance is to be practically applicable as a means of accurate behavioral prediction it is necessary that the process of intentional ascription be

corrigible in the light of accruing information about the behavior of the system and its environment. There is no means to this end other than the incorporation of witnessed behavior. It is interesting that Webb (1994) adopts something closely akin to reasoning that is at the heart of radical behaviorism here. Ideally, he argues, only behavior that has actually been witnessed should be considered, but this is impracticable. For instance, a person who has never smoked cigarettes before may write on a shopping list "Carton of cigarettes": the attribution of intentionality here cannot depend on whether someone else reads the shopping list. It is thus necessary to consider behavior that is only available in principle (Webb, 1994, p. 463). In radical behaviorist terms, the person making the list is a witness to his/her own behavior; it is immaterial that this opportunity is available only to him/her.

An important boundary condition that Webb introduces is that any behavior that the intentional stance considers must be displayed at the surface of the system that is being predicted (Webb, 1994, p. 464). Neuronal activity cannot be brought in to solve problems of the stance; only external behavior is permissible. Using information about brain states would make the intentional stance a subset of the design stance and make impossible the intended capacity to identify a pattern of behavior more abstract that that given by the design stance. Only behavior that systems display at their surface is thus admissible. This is witnessed behavior. Webb's initial plan is to

> . . .incorporate all behavioral evidence available in principle into the intentional stance. An intentional state is thus defined as whatever intentional state would be attributed on the basis of the intentional system's environment and exhibited behavior, where the "exhibited behavior" is to include all and only the behavior which (actually) occurred prior to its attribution." (p. 464)

The import of Webb's analysis is that Dennett has shown no means by which the intentional stance can attribute intentionality on the basis of witnessed behavior. Webb suggests that in incorporating some such method, Dennett would reveal himself as a behaviorist, a designation he does not necessarily deny but which he claims to abhor. None of this impinges on the argument being pursued here. My reading of Dennett fully accepts the necessity of attributing intentionality on the basis of witnessed behavior; Webb's analysis is a useful corroboration. Social cognitive psychology has evolved its own method of incorporating both the contextual and intentional stances and provides a means of overcoming the problems Webb raises, at least on a practical level. To this we now turn to lay the groundwork for two arguments: first, that the program of social cognitive psychology is incapable of implementation without the prior use of the contextual stance, and second, that the methods evolved by social cognitivists can be usefully employed to provide the necessary link between the extensional behavioral science and intentional cognitive psychology.

We need a system of behavior science that brings behavior in successfully (relating it to its environmental correlates) and avoids mereological problems. Hullian and Tolmanian systems will not do (Foxall, 2004); radical behaviorism is the sole survivor.

The Strategy of Intentional Behaviorism

What is Intentional Behaviorism?

The problem is not, however, the number of levels Dennett and his commentators posit. There is no single framework of analysis to be determined here but a series of pragmatic responses to an array of intellectual questions. Specifically, in the present context we are concerned with the adequacy of any of these levels to capture the essence of social cognitive psychology and to account for the phenomena it seeks to unravel. The import of the foregoing is that we must retain the sub-personal level of analysis and *for present purposes* demarcate its role as physiological, to which content is ascribed in order to enable prediction at the personal level. This is apparently Dennett's original position. The question arises as to the adequacy of this arrangement for the prediction and explanation of behavior.

The first difficulty it raises is the practical one of arguing from closely demarcated and localized physiological phenomena to mental constructs that are appropriate to the personal behavior of the organism. It has proved notoriously difficult for cognitive neuropsychology to isolate the neural substrates that can be convincingly correlated with closely-defined psychological or mental events other than those involved in basic sensory activity (Uttal, 2001). This is not a debate in which I wish to become embroiled since it is ultimately one that is open to specialists. I wish to note, however, the strong possibility that the kinds of neurocognitive inference Dennett wants to make may be considerably more controversial in practice than he assumes.

The second difficulty is that the sequence of logical reconstruction is, in any case, the reverse of what would be expected from the methods of modern cognitive neuroscience. In order to bring this into current focus, I should like to introduce the extended framework of analysis that I have called *intentional behaviorism*, which incorporates two bases of explanation.
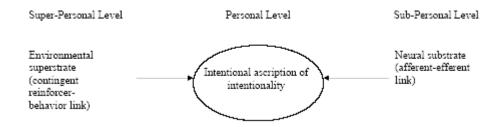

Fig. 1. Intentional Behaviorism.

The first, basically Dennett's (1969) approach, runs from the neural substrate of cognitive activity to the personal level of intentional ascription. Dennett's account begins, we have seen, with the neural event—specifically its role in an afferent–efferent process—and ascribes content on the basis of the resulting evolutionarily consistent logic. In other words, the direction of ascription is from neurology to intentionality, from the sub-personal level to the personal level of explanation. But this is not the actual procedure followed by neurophilosophy (Churchland, 1986, 2002), where the quest is for the physical substrate of behavior, including the expression of emotion, which is intentional. Take, for instance, the investigation of binocular rivalry. If two different and separate visual images are presented so that each can be seen by only one eye, they do not meld into a single image but compete for visual awareness. This phenomenon, in which each eye receives different inputs, is known as *binocular rivalry*. The resulting phenomenon is *bistable perception* in which, after a short period of confusion, the brain comes to receive the stimuli in an alternating sequence. Each stimulus is perceived for about one second, after which the other is received for the same interval, and so on, in an alternating fashion (Churchland, 2002, pp. 136-137.) This phenomenon can be investigated only if the experimental subject is conscious and able to report which of the images is currently perceived; there is no other means of establishing the neural correlates of the consciousness of each image.

Hence, the logical sequence of investigation is from the super-personal level to the sub-personal—from the verbal behavior of the participant to the physiological correlate(s) of both that behavior and any personal-level ascriptions of content that may be made. Intentional behaviorism requires that the procedure embrace the super-personal level of analysis, in which intentional ascription at the personal level is achieved via the observation of operant behavior (environment–behavior relationships) through extensional behavioral science. This is the second basis of explanation on which intentional behaviorism draws.

The purpose of the philosophical exercise that Dennett advances is, as he proposes, to ascertain what intentional content can be ascribed to the findings of neurological science, but the *de facto* procedure is more likely to entail using physiology and the logic of natural selection as a means of checking whether pre-ordained desires and beliefs can be rationally ascribed at the personal level. The consequent methodological procedure is thus: (1) the observation of environmental–behavioral relationships (including self-reports of emotion) at the super-personal level, leading to (2) the ascription of emotional content at the personal level, leading to (3) the search for the neural correlates of emotion at the sub-personal level. Desires and beliefs, and other mental content, are thus decided upon at the super-personal level as a result of the uncovering of environment–behavior links. Their appropriateness to this personal level ascription is further confirmed, however, by reference to the degree to which they can provide an additional heuristic overlay to the theories and findings of neurocognitive research at the sub-personal level on the basis of evolutionarily consistent reasoning.

Evidence of neural substrates of cognition (e.g., from fMRI scans) can show areas of the brain associated with mental activity such as thinking and emoting.

However, they cannot reveal the *content* of these mental events. This can be done only by probing the environment–behavior superstrates of cognition (e.g., by using the contextual stance). Therefore, Dennett's strategy in *Content and Consciousness* (1969) logically requires the further incorporation of the super-personal level of explanation through which confirmation of that the appropriate content is being ascribed. This requirement is doubtless implicit in his description of his strategy, but if his logical argument is to be completed it needs to be made explicit in terms of an extensional behavioral science based on the contextual stance. The superordinate framework of conceptualization and analysis presented by intentional behaviorism incorporates both Dennettian and Skinnerian analyses within a single, comprehensive account.

Strictly speaking, the contextual stance need not be restricted to operant psychology. It is simply a means of suggesting environment–behavior relationships that are consistent with selection-by-consequences and which can, therefore, act as indicators of the intentional content to be ascribed at the personal level.

Implementation of the Strategy

The strategy that Dennett advocates for the addition of content to physiological research may be followed in the case of operant behavioral science in order to generate a psychology of the person that takes environment–behavior relationships into consideration. The question arises: on what basis is content to be ascribed to theories and findings at the super-personal level in order to arrive at a psychology of the person that takes environment–behavior relationships into consideration? In order to find an answer to this question it is necessary to go back to Dennett's strategy of ascribing content to the sub-personal theories and findings of neuroscience, and it may be worthwhile reviewing its central themes now. At the same time, if the analogy between a sub-personal/personal-level linkage and a super-personal/personal level linkage is to be confirmed, it should be possible to show how the reasoning that develops for adding content to the extensional findings on environment–behavior relationships applies to the resolution of the problems of personal level psychology, behavioral continuity, and delimitation.

The ascription of content to the theories and findings of extensional behavioral science cannot be pursued in the absence of a convincing rationale. Recall that Dennett's strategy is to assume that the sequence of events that are to be intentionally explained are appropriate from an evolutionary perspective; the next step is to propose structures that will account for these appropriate sequences. The environmental significance necessary for the brain (in Dennett's terms) to discriminate useful from non-useful neural events is extrinsic to those neural events; the brain's necessary distinctions cannot stem solely from extensional descriptions of extrinsic stimulation and past behavior. The brain has to be able to discriminate and store fortuitously appropriate structures. Some close analogy of natural selection must be sought to provide for the capacity of the brain to do this. The necessary capacity could itself be an outcome of the evolution of species. An

intentional system has to be able to discriminate and respond to the environmental factors that impinge upon it, and to do this it must be able to "interpret peripheral stimulation." This entails producing within itself not representations but states or events that "co-occur" with the conditions or objects in its perceptual field. Information abstracted from the environment will nevertheless remain non-intelligent unless something else is added to it; what must be added consists in the detection of afferent and efferent links.

The links between the sub-personal/personal and super-personal/personal levels of analysis can, in each case, be characterized in Skinner's (1981) term "selection by consequences." The first is dependent on an evolutionary history that produced phylogenic consequences which determined the structure of the brain and its functioning, the neural afferent–efferent relationships to which content is added in the process of intentional ascription in order to delineate the personal level of analysis. The second also depends, indirectly, on this process since it is through natural selection that the organism's capacity to change as a result of contact with environmental consequences presumably came about. However, in a more direct way, this link is the result of ontogenic consequences through which behavior is shaped in the course of a lifetime. Again there is a need for intentional ascription, even if (or possibly, especially if) operant behavior instantiates physiological change within the organism. Donahoe, Palmer, and Burgos (1997, p. 196) state that "In a stable context, control by consequences (as opposed to antecedents) stands as a behavioral law, but we propose (at another level of analysis) that the effects of those consequences are implemented by changes in synaptic efficacies," an idea they trace back to Watson. But this argument merely addresses the sub-personal–personal levels of linkage that Dennett proposes and has no direct bearing on the relationship between the super-personal/personal levels, which are proposed here as a function of ontogenic development.

Intentional Ascription Revisited

We have seen that an extensionally-based system of radical behaviorist interpretation attempts to account for these necessary linkages by resorting to (a) physiological mechanisms, (b) private events, and (c) rules, and that there is no reason for taking any of these seriously at the explanatory level since they do not provide the necessary continuity, even in the terms required by an extensional science of behavior. The required interpretative device is that of content ascription in terms of the desires and beliefs it would be rational for the individual to have in view of his or her situation defined by the intersection of his or her learning history and the behavior setting he or she faces. Both evolutionary reasoning and the behavioral analysis of matching phenomena suggest that the contingencies with which an individual will have come most obviously into contact in the course of phylogenic and ontogenic histories will be those producing behavior that tends toward optimization of outcome. In any situation, therefore, we can assume beliefs, attitudes, and intentions that are consistent with this objective. As long as the conceptualization and—at the level of empirical research—measurement of these

cognitive constructs is in line with those pursued by attitude theorists, there is a convincing rationale for the attribution of content to the findings of extensional behavioral science based on the contextual stance (i.e., the location of behavior at the intersection of learning history and behavior setting). These constructs directly link the elements of the contextual stance with the process of content ascription.

The import of intentional ascription must, however (in the course of the present argument) go thus far and no further. It retains the a-ontological assumption about cognitive events, states, processes, and structures with which, along with Dennett, we began. There is no justification for uncritically accepting the entire apparatus of the information processing account of behavior whether based on cognitive conjecture or neurophysiology. The justification of intentional behaviorism lies in the necessity of connecting efferent–afferent processes in some way that (a) physiology cannot, (b) behavioral science cannot, and (c) that aids the coherent explanation and prediction of behavior. What Dennett calls a centralist theory, therefore, has two explanatory components. The first is an extensional account of the interaction of functional structures. The second is an intentional characterization of these structures, the events occurring within them, and states of the structures resulting from them. The links between the extensional account and the intentional interpretation consists of a hypothesis or hypotheses describing the evolutionary source of the fortuitously propitious arrangement in virtue of which the system's operation in this instance makes sense. These hypotheses are required, in principle, to account for the appropriateness which is presupposed by the intentional interpretation but which requires a genealogy from the standpoint of the extensional, physical theory. Despite the inevitable imprecision of this approach, the challenge is to make the case that the ascription of content to the theories and findings of behavioral science can be of use to the behavior analyst, and in particular to the process of radical behaviorist interpretation.

As Gunderson (1972) summarizes Dennett's argument, humans are not simply neurophysiological organisms but also persons who exhibit complex behaviors. Dennett's case for the ascription of content rests on the understanding that because some neural events, states, and structures are about other things (i.e., intentional) it is possible to ascribe content to them. The basis of the contextual stance is similarly that humans are persons as well as organisms whose behavior is determined by the contingencies of reinforcement. Moreover, some of the environmental elements on which our behavior is contingent are *about* things (i.e., are such that it makes sense to attribute content to them, to add an extra layer of interpretation that is relevant to the personal level). While Dennett speaks of only two levels of analysis, however, we have distinguished three. We have noted his argument for a *personal* level, at which the individual as a whole discriminates such "mental" entities as pain, and a *sub-personal* level of brains and neurons, at which level the physiological correlates of pain behavior can be detected:

> . . .[T]he terms in our mentalistic vocabulary are nonreferring. Rather like "sakes" or "miles", [or centers of gravity] mentalistic terms in appropriate contexts tell us something, but succeed in doing so without thereby referring to

any entities any more than the words "sakes" or "miles" refer to sakes or miles. (Gunderson, 1972, p. 593)

At the super-personal level we turn to the environmental contingencies that shape and maintain responding in order to find an extensional basis for the ascription of such content. Several factors distinguish this level from both the personal and the sub-personal level based on neuroscience that Dennett identifies.

First, the super-personal level cannot capture anything of the personal level, including some essential components of what it is to be human such as being able to discriminate pain. No matter how we grimace and howl and hold our painful heads, no matter what consequences these overt actions have by way of producing sympathy or medicine or exemptions from work from others, these super-level events are entirely separate from the discrimination of pain. Second, the super-personal level constitutes an extensional approach to the science of behavior, one which can explain much behavior at that level but which is incapable of dealing with the things that can only be discriminated at the personal level such as pain, that it is time to go home, and other intentional matters. Only by the addition of a heuristic overlay of interpretation can these personal-level matters be accommodated. Third, even though neither level reduces to the other, it is incumbent upon us to show how they are linked if we are to make legitimate and convincing interpretive ascriptions. The link, moreover, must be consistent with evolutionary reasoning. There are several strands to be considered here:

(a) The capacity for operant reinforcement is bestowed by natural selection. What Skinner (1981) calls "selection by consequences" is the analogy/homology that links the two processes at least at the level of phylogenic and ontogenic consistency.

(b) In the case of linking the personal and sub-personal levels, the links must supervene (i.e., add appropriate interpretation) between the afferent and efferent processes of the brain. The corresponding processes in operant conditioning are stimuli and responses. The heuristic overlay of intentionality must link these in ways that an extensional account cannot. There are three such ways: (i) to elucidate the personal level, (ii) to demonstrate continuity of behavior from setting to setting, and (iii) to solve problems of equifinality by delimiting operant interpretations that (attempt to) proceed solely at the extensional level.

These considerations bring the interpretation within the scope of an evolutionarily consistent framework of conceptualization and analysis. How? The animal that is to be successful in negotiating its environment must be differentially sensitive to discriminative and other setting stimuli in order to act appropriately (with behavior that will be reinforced).

There is no more reason to believe that a physiological account will eventually be available to show how this occurs any more than there is a possibility that a physiological account will be able to demonstrate an individual's discrimination of pain. The discriminative control of appropriate behavior occurs at the personal level. The recognition of appropriate inaugurating stimuli is a similar process. At the very least, the intentional mode of explanation cannot be

abandoned until the physiological link is demonstrated. To trust in eventual physiology is superstitious in a way in which the ascription of intentionality is not (if the latter strategy results in more effective predictions of behavior). Physicists who shun the concept of center of mass in favor of a belief in some more remote physical explanation would be showing a similar level of superstition. That physicists are not embarrassed to include centers of mass in their predictive work should be an example to the psychologist.

Intentional Behaviorist Interpretation

The alternative to an extensional system of radical behaviorist interpretation, then, is the amalgamation of extensional operant behavioral science and Dennett's intensional stance, by which content would be ascribed to its theories and findings in order to provide a basis for radical behaviorist interpretation. The reality of this may be closer to us than we have imagined. The point is sometimes made that radical behaviorists often incorporate the language of intentionality in their popular accounts of behavior, the implication being that the extensional operant account is thereby diminished and perhaps incapable of adequately describing the events that are the subject of the accounts in question. Skinner (e.g., 1974) argues that in order to communicate to a non-specialist audience it is useful to adopt everyday language, as does the professional astronomer speaks of the sun "rising" and "setting" when addressing children. Many behaviorists have taken this at face value and not concerned themselves further with the charge that the use of such language necessarily invokes a theoretical stance that is inevitable in the explanation of behavior. In view of the import of the current argument, this is a serious matter that behavior analysts ought not to ignore so easily.

The accounts in question are generally interpretations rather than reports of experimental work, and this suggests that at least at the level of interpretation intentional language is inevitable not only to communicate to pedestrians but to express the ideas involved in accounting for complex activity in operant terms. "Thinking" and "feeling," the very stuff of private events, are almost always spoken of in intentional language: we do not just think, we think *about* or think *that*; we do not just feel, we feel *that*; and so on. We can treat such events as stimuli and responses that do not differ in kind from those that are publicly available, though this is to make an enormous ontological leap that can never be the subject of a scientific analysis—but to insist that thoughts and feelings are simply discriminative stimuli (or establishing operations, or other source of antecedent stimulation), associating them in the process with a physiological level of extensional analysis, is to leave out entirely the personal level to which Dennett draws attention, the level without which no psychological explanation can be complete.

The suggested program is not a call for the use of mediating events or the kinds of theory that Skinner repudiated. Even less is it a regurgitation of the sometimes argued notion that the intentional and contextual stances might be conjoined or a synthesis generated that would combine "the best of each." This is

not possible in practice because their respective intentional and extensional bases are incommensurable (Foxall, 1999). But the adding of content to an extensional account is not a synthesis or amalgamation. It is not adding anything to the findings and theories derived from the experimental analysis of behavior. Rather, it is the derivation of another level of interpretation in order to facilitate understanding and prediction by taking the personal level of experience into account.

In order to advance the debate between cognitivists and behaviorists, this account takes Dennett's thesis about the relationship between extensional science and intensional psychology at face value. To do this is to share, again for the sake of argument, (a) his assessment of the (literal) shortcomings of purely extensional science as a means to understand behavior: such science simply does not go far enough in the quest to explain all behavior, and (b) his judgment that the link between the two is found in the imperatives of behavioral science is, like physiology, an autonomous approach to knowledge in its own right, but it is incapable of explaining all human behavior within its own theoretical and methodological purview. Nor can it even engender plausible interpretations (expressed in non-convoluted language) of all behavior. Here is an important parallel with Dennett's analysis leading to a major conclusion: the extensional science of physiology is to Dennett's intensional physical psychology what an extensional behavioral science is to the intensional psychology of social cognition. In other words, the extensional science provides the evolutionary basis for understanding behavior biologically to which intensional cognitive interpretation verbally ascribes an a-ontological, initially non-empirical dimension which yields predictions of certain behaviors that the extensional approach (of itself) can neither explain nor predict. What is true for the centerpiece of social cognitive psychology—attitude research—is likely to be generally the case.

The strategy of ascribing optimality (rationality) to systems in order to predict their behavior is a methodological simplification that involves further ascription—of posited entities such as beliefs, attitudes, and intentions which, as we have seen, have the function of fine-tuning the prediction by linking it to the system's environmental history and behavior setting. The three stages of the intentional strategy make its dependency on the prior application of the contextual strategy clear. Dennett takes pains to avoid this conclusion. He denigrates (radical) behaviorism by, firstly, casting it as a simplistic S-R paradigm, and, secondly, by asserting, in the absence of any adduced evidence, that it has proved to be unsuccessful in predicting behavior. The first of these caricatures fails to engage with the operant behavior analysis of the last thirty years, especially the analysis of behavior at the molar level, the post-Skinnerian analysis of the verbal behavior of the listener, etc. The second ignores a mass of empirical evidence, and both overlook the possibility of radical behaviorist interpretation, that is, the use of the contextual stance to account for the behavior that is not amenable to an experimental analysis. Indeed, the use of the intentional stance is advocated here only in the context of radical behaviorist interpretation. It is important that the extensional science of operant behavior analysis continue its program for two

reasons: firstly, to provide an evolving and expanding base for the content ascription to which content can be ascribed in the process of interpretation; and secondly to provide alternative, competing, and challenging explanations. Insofar as the growth of knowledge depends on "the active interplay of competing theories" (Feyerabend, 1975), it is essential to have (i) a thriving experimental analysis of behavior, (ii) operant interpretations which themselves attempt to function on an extensional level only, and (iii) operant interpretations that contain the intentional overlays necessary to provide accounts of behavior at the personal level. Their interaction is, indeed, a *sine qua non* of intellectual progress. Hence, what characterizes the intentional behaviorist approach is the incorporation of both the contextual and the intentional stances into a single framework of analysis. Social cognitivists must reconstruct desires and beliefs in the context of the individual's rationality by considering its situation. The contextual stance facilitates this reconstruction by *de*constructing the notion of situation in terms of (a) a learning history, (b) the current behavior setting, and (c) their interaction. This is both consistent with and a means of operationalizing Dennett's view that the organism will have those desires and beliefs that are appropriate to it given its situation.

Intentional behaviorism differs from the other systems of explanation in its comprehensive inclusion of the various elements of the contextual and intentional stances as well as in the understanding that the ascription of intentionality reinforces rather than detracts from the prior existence of an extensional behavioral science. It follows Dennett's subtle recognition that the addition of an intentional layer of interpretation does not discover anything new but tells another story about the theories and findings produced by operant psychology. The result is not just an extra story that maps on to the original in a one on one fashion; rather, it extends the scope and relevance of the interpretation. Moreover, intentional behaviorism recognizes that social cognitive psychology proceeds in a similar manner and it raises the possibility that psychology will find a platform on which it might unite.

The Strategy of Super-Personal Cognitive Psychology

What is Super-Personal Cognitive Psychology?

Another kind of contextual psychology, super-personal cognitive psychology, represents the refinement of the intentional terms employed in intentional behaviorism so that they can enter into scientific theories that employ the contextual stance. This is a performance theory, but the questions remain (a) of the ontological status of the entities proposed, and (b) of their causal significance. There is no need to accord them either table-and-chair reality (though they may be real in Dennett's sense that parallelograms of forces are "real") or causal efficacy unless this is derived from molar patterns of environment–behavior relationships in which the contingencies can be shown to control behavior. There is limited acceptance that humans can be creative in formulating personal or self-rules, but there is not (and cannot be) any evidence that these procedures are not

environmentally determined through environmental–behavioral conditioning. The link with evolution may come in this case through evolutionary psychology. The reason is that we are now seeking cognitive ascriptions that can be justified directly in terms of evolutionary processes. Evolutionary psychology is concerned with the development of cognitive structures that were appropriate during the Pleistocene period and which are an amalgam of biological and mental/behavioral selection. Moreover, it is possible for super-personal cognitive psychology to point the way toward a comprehensive cognitive–behavioral psychology.

Theories that deal only in identifying the necessary intentional idioms to explain behavior (e.g., the Bolles-Bindra-Toates expectancy theory) function at the level of intentional behaviorism. They are competence models, and as such they seek semantic understanding. A fully realized super-personal cognitive psychology, like Dennett's sub-personal cognitive psychology, has the potential to uncover the syntax of the brain but through (unlike Dennett) relating environment–behavior consistencies to physiological structures and functions. One familiar component of this project is the attempt to relate environment–behavior relationships to dopamine release. If this can be shown to be evolutionarily consistent, it justifies the ascription of the appropriate intentional content at the personal level. Similarly, research shows that pain is a personal-level phenomenon that can be systematically related to operant conditioning (Flor, Knost, & Birbaumer, 2002) as well as neural substrates. This is more than establishing that operant conditioning leads to dopamine release, or the differential locations of neurons that fire when alternative brands of soft drinks are presented, or when different kinds of decision making are occurring—though all of these are relevant. It means establishing relationships between sequences of environmentally-maintained behavior (operant conditioning) and neurological activity, a link between the super-personal and the sub-personal, from the whole complex of which an inference can be made about what the organism is doing at the personal level. Pain, for instance, would not be inferred from either a sub-personal physiological pattern of afferent–efferent linkages or a super-personal relationship between certain verbal and nonverbal behaviors like screaming or holding one's thumb after hitting it with a hammer, but from evidence that connected the two. Flor and her colleagues (Flor, Knost & Birbaumer, 2002) have demonstrated, for instance, that the reports of pain made by back pain sufferers are susceptible to operant control. The solicitous behavior of a spouse, or their mere presence in the same room, provides reinforcement of the pain sufferer's verbal reports of pain; brain wave activity associated with pain has been identified only when the spouse was present to complete the operant training.

Implementation of the Strategy

The tentative strategy of super-personal cognitive psychology outlined here has four stages. It begins with the intentional idioms found to be useful at the level of intentional behaviorism, but, additionally (and wherever possible), those that have been supported by empirical testing. An example of the intentional

behaviorism phase of this stage is incorporation of the S-S* and R-S* links proposed by Bolles's syllogism. Dickinson's (1997) program to substantiate these proposed links comprises an empirical program. Indeed, in the absence of empirical evidence, at least for epistemological usefulness of the content introduced in the process of intentional behaviorism, one would have to rethink the program before proceeding to this stage. There is room, however, for theoretical reasoning as to the mental processes that would be required to accomplish pre-behavioral information processing.

The second stage is to employ the psychological structures further by enquiring how they are implemented in the brain. The necessary sequence in super-personal cognitive psychology is that followed by Dickinson:

> Whereas in the theoretical vacuum of the 1960s I had hoped that the pattern of behavioral dysfunction produced by neural interventions would reveal the psychological structures and processes of instrumental action, I now have sufficient confidence in the present psychological understanding to reverse the research strategy and ask how the brain implements these processes and structures. (Dickinson, 1997, p. 361)

The third stage involves relating the brain processes to specific afferent–efferent linkages and their evolution in the process of natural selection. This reinforces the original ascription of intentionality made in intentional behaviorism at the personal level. That level is, of course, where they remain.

Finally, it is necessary to relate the behavior–environment relationships on the basis of which the intentional behaviorism ascriptions were made to the evolution of such behaviors in the process of evolutionary psychology. This procedure is not a prescribed route for empirical science but reflects how some scientists have worked and how, as a result, science has progressed in the direction of what I have called super-personal cognitive psychology. Above all, it is iterative and corrigible.

Summing Up

The difference between intentionalistic explanation and that which characterizes radical behaviorism is a difference between the sentences employed to express them. Specifically, the propositions in terms of which intentionalistic explanation is expressed are not substitutable for other propositions even though they are equivalent in meaning. The essence of radical behaviorism as a philosophy of psychology is the explanation of behavior without recourse to propositional attitudes, without attributing propositional content to its statements of behavior–environment relations. Hence the components of the patterns of contingency that provide its fundamental explicatory device are all defined in *ex*tensional terms, be they establishing operations, discriminative stimuli, responses, or positive/aversive consequences. Private events as well as overt behaviors and verbal as well as non-verbal behaviors are all defined in terms that exclude the intentional idioms. The extensional language, which is generally taken to be the only form that scientific discourse ought properly to take, is said to be

"transparent" in that coextensive terms can be substituted one for another. We can equally say that the animal's behavior is reinforced by the food contingent on it and that it is reinforced by the protein pellet that is presented after each response. "Food" and "protein pellet" have the same extension in this context.

Radical behaviorism's *explanation* of behavior is deficient in three ways. These are particularly evident in its attempt to interpret complex behavior (i.e., behavior not amendable to direct experimental analysis) by reference to behavioral principles established through the basic and applied analysis of behavior conducted in the laboratory and the field. I do not mean by this that radical behaviorism fails to predict behavior or to make possible its influence or control. I do not, therefore, deny that it identifies environmental events that are the independent variables of which behavior is a function. Indeed, I have no criticism of radical behaviorism in terms of its own pragmatic success criterion, the prediction (often plausible postdiction) and control (influence) of behavior. Here it undoubtedly succeeds, not only in the closed settings of the operant laboratory and therapeutic community but in the open settings presented by the economics of everyday life. Nor am I unaware that for many behavior analysts this is sufficient to constitute an explanation of behavior, as it is in the Machian positivism that guided Skinner's early construction of this paradigm and from which he never departed. In other words, I am not criticizing the capacity of radical behaviorism to succeed on its own terms. I *am* arguing that on those terms radical behaviorism cannot account for (a) behavior at the personal level (as opposed to accounting for behavior–environment relationships), (b) the continuity of behavior over time and space, and (c) the delimitation of interpretations of behaviors. The picture is complicated by dint of the insistence of radical behaviorists that these putative problems can be overcome by means of private events, verbal behavior, rule-governance, relational frames, and other devices. I argue that none of these, in fact, provides the necessary conceptual framework for a science of behavior that succeeds in accounting for the three requirements in which I find it deficient.

Hence, I am not advocating that radical behaviorism as it stands is in need of radical change. It is as important to the program I advocate as it is to its own that it continue to develop as a science of behavior that elucidates behavior–environment relationships. It is a means of predicting and influencing behavior, one of several psychologies that attempt these tasks, and it is important that it continue to provide a limiting statement of how behavior may be determined by its consequences. But it is equally important that it be challenged by a theoretical approach that identifies its deficiencies in the realm of explanation into providing a robust experimental program charged with showing that such empirical work can provide a full explanation of behavior. I *do* argue, however, that explanation which deals with the personal level of explanation, which accounts for continuity, and which is delimited requires that subtle use be made of the intentional psychology that underlies cognitivism. This takes nothing away from radical behaviorism; it adds what is necessary to explain and interpret rather than simply to predict and control.

The paper has not argued against any philosophy of psychology; rather it has sought to propose a synthetic framework within which each may contribute to the

explanation of complex human behavior. Each is essential to the research program of intentional behaviorism in its own right. It is essential, for instance, that research within the extensional behavioral science context of radical behaviorism continues, that the possibility of accounting for behavior in its terms alone be explored comprehensively. It is equally important that a cognitive program proceed in its own right. It is also vital that they interact tenaciously, for it is the competitive collisions of the explanations produced by these schools of thought as well as the development within each of their confines that engender the progress of science.

References

Angulo, M. C., Staiger, J. F., Rossier, J., & Audinat, E. (1999). Developmental synaptic changes increase the range of integrative capabilities of an identified neocortical connection. *Journal of Neuroscience*, *19*, 1566-1576.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Baum, W. (1973). The correlation based law of effect. *Journal of the Experimental Analysis of Behavior*, *20*, 237-253.

Baum, W. M., & Heath, J. L. (1982). Behavioral explanations and intentional explanations in psychology. *American Psychologist*, *47*, 1312-1317.

Bechtel, W. (1988). *Philosophy of mind*: *An overview for cognitive science*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Oxford: Blackwell.

Berridge, K. C. (2000). Reward learning: Reinforcement, incentives, and expectations. In: D. L. Medin (Ed.), *The psychology of learning and motivation, 49* (pp. 223-278). San Diego: Academic Press.

Bindra, D. (1978). How adaptive behavior is produced: A perceptual-motivation alternative to response reinforcement. *Psychological Review*, *81*, 199-213.

Bolles, R. C. (1972). Reinforcement, expectancy, and learning. *Psychological Review*, *79*, 394-409.

Bouton, M. E., & Franselow, M. S. (Eds.) (1997). Learning, motivation, and cognition: The functional behaviorism of Robert C. Bolles. Washington, DC: American Psychological Association.

Catania, A. C. (1992). *Learning* (3rd ed). Englewood Cliffs, NJ: Prentice-Hall.

Chisholm, R. (1957). *Perceiving: A philosophical study*. Ithaca, NY: Cornell University Press.

Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind/brain*. Cambridge, MA: MIT Press.

Churchland, P. S. (2002). *Brain-wise: Studies in neurophilosophy*. Cambridge, MA: MIT Press.

Davidson, D. (1980). *Essays on action and events*. Oxford: Clarendon Press.

Davies, M. (2000). Persons and their underpinnings. *Philosophical Explorations*, *3*, 42-62.

Dennett, D. C. (1969). *Content and consciousness*. London: Routledge.

Dennett, D. C. (1976). Conditions of personhood. In: A. Rorty (Ed.), *The identities of persons* (pp. 175-196). Berkeley, CA: University of California Press.

Dennett, D. C. (1978). *Brainstorms*. Montgomery, VT: Bradford.

Dennett, D. C. (1981). Three kinds of intentional psychology. In: R. Healy (Ed.), *Reduction, time, and reality* (pp. 37-61). Cambridge: Cambridge University Press.

Dennett, D. C. (1982). How to study human consciousness empirically, or nothing comes to mind. *Synthese*, *59*, 159-180.

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *The Behavioral and Brain Sciences*, *6*, 343-390.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (1991). *Consciousness explained.* Boston, MA: Little, Brown.

Dickinson, A. (1997). Bolles's psychological syllogism. In: M. E. Bouton & M. S. Fanselow (Eds.), *Learning, motivation, and cognition: The functional behaviorism of Robert C. Bolles* (pp. 345-268). Washington, DC: American Psychological Association.

Donahoe, J. W., Palmer, D. C., & Burgos, J. E. (1997). The S-R issue: Its status in behavior analysis and in Donahoe and Palmer's *Learning and complex behavior. Journal of the Experimental Analysis of Behavior*, *67*, 193-211.

Elton, M. (2000). Consciousness: Only at the personal level. *Philosophical Explorations*, *3*, 25-41.

Elton, M. (2003). *Daniel Dennett: Reconciling science and our self-conception*. Cambridge: Polity Press.

Feyerabend, P. (1975). *Against method*. London: NLB.

Flor, H., Knost, B., & Birbaumer, N. (2002). The role of operant conditioning in chronic pain: An experimental investigation. *Pain*, *95*, 111-118.

Foxall, G. R. (1997). Explaining consumer behavior: From social cognition to environmental control. In: C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology*, *12* (pp. 229-287). Chichester: Wiley.

Foxall, G. R. (1999). The contextual stance. *Philosophical Psychology*, *12*, 25-52.

Foxall, G. R. (2004). *Context and cognition: Interpreting complex behavior.* Reno, NV: Context Press.

Foxall, G. R. (2007a). *Explaining consumer choice.* New York: Palgrave Macmillan.

Foxall, G. R. (2007b). Explaining consumer choice: Coming to terms with intentionality. *Behavioral Processes*, *75*, 129-145.

Gardner, S. (2000). Psychoanalysis and the personal/sub-personal distinction. *Philosophical Explorations*, *3*, 96-119.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (1998). *Cognitive neuroscience: The biology of the mind.* New York: W. W. Norton.

Gunderson, K. (1972). Content and consciousness and the mind-body problem. *The Journal of Philosophy*, *69*, 591-604.

Hayes, S. C., Zettle, R. D., & Rosenfarb, I. (1989). Rule-following. In: S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 191-220). New York: Plenum.

Hempel, C. G. (1980). The logical analysis of psychology. In: N. Block (Ed.), *Readings in philosophical psychology*, Vol. I (pp. 14-23). London: Methuen.

Herrnstein, R. J. (1997). *The matching law: Papers in psychology and economics*. (H. Rachlin & D. I. Laibson, Eds.) New York: Russell Sage Foundation; Cambridge, MA: Harvard University Press.

Heyes, C. (2000). Evolutionary psychology in the round. In: C. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 3-22). Cambridge, MA: MIT Press.

Hornsby, J. (2000). Personal and sub-personal: A defence of Dennett's early distinction. *Philosophical Explorations*, *3*, 6-24.

Juarrero, A. (1999). *Dynamics in action: Intentional behavior as a complex system.* Cambridge, MA: MIT Press.

Kandel, E. R. (2001). The molecular biology of memory storage: A dialogue between genes and synapses. *Science*, *294*, 1030-1038.

Lacey, H. M. (1995/1996). Behaviorisms: Theoretical and teleological: A review of John Staddon's *Behaviorism: Mind, mechanism, and society*, and Rachlin's *Behavior and mind*: *The roots of modern psychology. Behavior and Philosophy*, *23*, 61-78.

Lee, V. L. (1988). *Beyond behaviorism.* London: Erlbaum.

Mach, E. (1896/1959). *The analysis of sensations and the relation of the physical to the psychical* [C. M. Williams, Trans.]. New York: Dover.

Mach, E. (1905/1976). *Knowledge and error: Sketches on the psychology of enquiry* [T. J. McCormack & P. Foulkes, Trans.]. Boston, MA: Reidel.

Malcolm, N. (1977). Behaviorism as a philosophy of psychology. In: N. Malcolm (Ed.), *Thought and knowledge* (pp. 85-103). Ithaca, NY: Cornell University Press. First published in T. W. Wann (Ed.) (1964), *Behaviorism and phenomenology: Contrasting bases for modern psychology* (pp. 141-162). Chicago, IL: University of Chicago Press.

McDowell, J. (1994). *Mind and world.* Cambridge, MA: Harvard University Press.

McGinn, C. (1991). *The problem of consciousness*. Oxford: Blackwell.

McGinn, C. (2004). *Consciousness and its objects.* Oxford: Oxford University Press.

Moore, J. (1999). The basic principles of behaviorism. In: B. Thyler (Ed.), *The philosophical legacy of behaviorism* (pp. 41-68). Dordrecht: Kluwer Academic Publishers.

Morse, W. H., & Kelleher, R. T. (1977). Determinants of reinforcement and punishment. In: W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 174-200)*.* Englewood Cliffs, NJ: Prentice-Hall.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Rachlin, H. (1989). *Judgment, decision, and choice: A cognitive/behavioral synthesis*. San Francisco, CA: Freeman.

Rachlin, H. (1994). *Behavior and mind: The roots of modern psychology*. New York: Oxford University Press.

Rachlin, H. (2000). Teleological behaviorism. In: W. O'Donohue & R. Kitchener (Eds.), *Handbook of behaviorism* (pp. 195-215). San Diego: Academic Press.

Schnaitter, R. (1999). Some criticisms of behaviorism. In: B. A. Thyer (Ed.), *The philosophical legacy of behaviorism* (pp. 209-249). Dordrecht: Kluwer.

Searle, J. (1983). *Intentionality*: *An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*, *52*, 270-277.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

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

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

Skinner, B. F. (1981). Selection by consequences. *Science*, *213*, 501-504.

Smith, L. D. (1986). *Behaviorism and logical positivism: A reassessment of the alliance*. Stanford, CA: Stanford University Press.

Smith, T. L. (1994). *Behavior and its causes: Philosophical foundations of operant psychology*. Dordrecht: Kluwer.

Staddon, J. E. R., & Cerutti, D. T. (2003). Operant conditioning. *Annual Review of Psychology*, *54*, 115-144.

Strawson, G. (1994). *Mental reality.* Cambridge, MA: MIT Press.

Toates, F. (1986). *Motivational systems.* Cambridge: Cambridge University Press.

Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain.* Cambridge, MA: MIT Press.

Webb, S. (1994). Witnessed behavior and Dennett's intentional stance. *Philosophical Topics*, *22*, 457-470.

Zettle, R. D., & Hayes, S. C. (1982). Rule-governed behavior: A potential framework for cognitive–behavioral therapy. In: P. C. Kendall (Ed.), *Advances in cognitive–behavioral research and therapy* (pp. 73-117). New York: Academic Press.