

# A CRITICAL EXAMINATION OF B. F. SKINNER'S OBJECTIONS TO MENTALISM<sup>\*1</sup>

Russell Keat<sup>+</sup>

## I

One of the most frequent objections to behaviorism in psychology has been that it either denies the existence of, or at least fails to deal adequately with, man's "inner life." A typical example of this kind of objection is contained in the following passage from Carl Rogers' contribution to a symposium on "Behaviorism and Phenomenology":

"Valuable as have been the contributions of behaviorism, I believe time will indicate the unfortunate effects of the bounds it has tended to impose. To limit oneself to consideration of externally observable behavior, to rule out consideration of the whole universe of inner meanings, of purposes, of the inner flow of experiencing, seems to me to be closing our eyes to great areas which confront us when we look at the human world." ([40], p. 119).

To those who find this objection a powerful one, it must often seem that the form of behaviorism espoused by B.F. Skinner is quite unacceptable. For, quite apart from the general impression one may gain that Skinner's behaviorism is in some way "extreme," his work contains a sustained and varied polemic against what he terms "mentalism," i.e., the employment of mentalistic concepts in the analysis and explanation of behavior. Indeed, the polemic has become even more fervent in his recent book, *Beyond Freedom and Dignity*, since Skinner clearly regards mentalistic approaches as one of the main stumbling blocks to an adequate "technology of human behavior," a technology which is urgently required to solve the enormous social problems that increasingly threaten the survival of the human species. A mentalistic psychology encourages us to adopt a systematically mistaken approach to those problems:

---

\* [2013] Published in *Behaviorism*, 1 (1), 1972, 53-70: citations should be to this. The analysis of Skinner's work presented here was informed by extensive conversations with (the late) Professor Willard F Day, whom it was my great good fortune to have met, and become friends with, when I held a visiting position at the Department of Psychology, University of Nevada at Reno, in 1969.

+ [2013] School of Social and Political Science, University of Edinburgh: russell.keat@ed.ac.uk. Previously, Department of Philosophy, Lancaster University.

“Every issue of a daily paper, every magazine, every professional journal, every book with any bearing whatsoever on human behavior will supply examples. We are told that to control the number of people in the world we need to change *attitudes* toward children, overcome *pride* in size of family or in sexual potency, build some *sense of responsibility* toward offspring, and reduce the role played by a large family in allaying *concern* for old age. To work for peace we must deal with the *will to power* or the *paranoid delusions* of leaders; we must remember that wars begin in the *minds* of men, that there is something suicidal in man — a *death instinct* perhaps — which leads to war, and that man is aggressive by *nature*. To solve the problems of the poor we must inspire *self-respect*, encourage *initiative*, and reduce *frustration*. To allay the disaffection of the young we must provide a *sense of purpose* and reduce feelings of *alienation* or *hopelessness*. Realizing that we have no effective means of doing any of this, we ourselves may experience a *crisis of belief* or a *loss of confidence*, which can be corrected only by returning to a *faith in man’s inner capacities*. This is staple fare. Almost no one questions it.” ([39], pp. 9-10).

Yet despite this vigorous rejection of mentalism, it is by no means clear that the “inner life” objection can be maintained against Skinner. For his “radical behaviorism”<sup>2</sup> pays considerable attention to what he terms “private events,” and though he resolutely affirms their strictly *physical* status, we find amongst them a large number of items that would normally be invoked by the defenders of man’s “inner life” - such as joy and sorrow, pain, and various “feelings” (see, e.g. [33], pp. 257-60). As early as the 1945 symposium on Operationism, Skinner explicitly contrasted his own “radical behaviorism” with the more conventional varieties, partly in terms of its inclusion of private events. The following passage makes this clear:

“The distinction between public and private is by no means the same as that between physical and mental. That is why methodological behaviorism (which adopts the first) is very different from radical behaviorism (which lops off the latter term in the second). The result is that while the radical behaviorist may in some cases consider private events (inferentially, perhaps, but none the less meaningfully), the Boring-Stevens operationist has maneuvered himself into a position where he cannot. ‘Science does not consider private data,’ says Boring. (Just where this leaves my contribution to the present symposium, I do not like to reflect.) But I contend that my toothache is just as physical as my typewriter, though not public, and I see no reason why an objective and operational science cannot consider the processes through which a vocabulary descriptive of a toothache is acquired and maintained. It is an amusing bit of irony that, while Boring must confine himself to an account of my external behavior, I am still reasonably interested in what might be called Boring-from-within.” ([32], p. 294).

And, in the subsequent *Science and Human Behavior*, a separate chapter, “Private Events in a Natural Science,” is devoted to this topic ([33], chapter XVII). Thus for Skinner, private events are regarded as theoretically acceptable components in the description of contingencies of reinforcement, despite the difficulties deriving from their relative inaccessibility to public (mainly verbal) control.<sup>3</sup>

But although some mentalistic items “re-emerge” in radical behaviorism as private events, many others fall by the wayside - or rather, they are given what is *apparently* a reductive “translation” into behavioral dispositions. Precisely how we are to regard these translations is a much-disputed issue, and will not be discussed here.<sup>4</sup> But examples of such “translatable” terms include “intelligence,” “greed,” and “ambition” (see, e.g. [38], p. 276). Unfortunately, Skinner never attempts to state explicitly what it is that distinguishes these examples from those mentioned earlier, i.e., “joy,” “sorrow,” “pain,” etc., but it seems that something like this is involved: terms referring to personality-traits, abilities and motives are given translations, whilst those with a strongly *experiential* reference remain as denoting private events.

In most of what follows, I shall be examining critically the various objections that Skinner has made against mentalism (noting, at some points, what seem to me certain changes in emphasis in different periods of his writings). Though most of the objections are closely interrelated, I shall for convenience adopt the following classification: for Skinner, mentalism is to be rejected because—

- (1) It lacks “explanatory power.”
- (2) It involves the employment of “theories.”
- (3) It tends to invoke “homunculi” or “inner agents”, regarding man as an “autonomous agent”.
- (4) It distracts our attention from the study of *behavior*.
- (5) It involves a dualistic ontology of the “mental” and the “physical”.

I shall argue that objections (1) - (4) are at best inconclusive, and that although he often denies attaching much significance to (5), this may well in fact be the most powerful objection. But in the final section, I shall return to some questions concerning the role of private events in radical behaviorism, especially this one: what is it about private events that makes them acceptable to Skinner, in contrast to their mentalistic counterparts? For it seems that they may be open to at least some of the objections to mentalism, and that their only “advantage” derives from their physical

status. Yet at the same time, Skinner seems unwilling to consider the philosophical problems involved in ascribing this status to them.

But it is not only at this point that Skinner's position raises philosophical issues: they are also relevant to many of the arguments against mentalism contained in objections (1) - (4). In particular questions concerning the nature of explanation, much discussed in the recent literature in the philosophy of science, are frequently involved, and one underlying purpose of this paper is to suggest that these distinctively philosophical issues cannot be avoided by any radical behaviorist.<sup>5</sup> It might seem to some that this claim hardly needs to be made explicit: but some interesting encounters with radical behaviorists, and more importantly, a certain element in Skinner's own thinking, have made me think otherwise. The element concerned is roughly this: the tendency to attempt a *replacement* of the evaluative-prescriptive mode of analysis by an empirical-descriptive one.

Thus, although Skinner would clearly agree that issues concerning explanation are central to the rejection of mentalism (see, e.g. [38] p. 222), it is quite possible that he would prefer to substitute, for a discussion concerning the legitimacy and philosophical acceptability of certain modes of explanation, an empirical investigation of the contingencies of reinforcement involved in the emission of verbal behaviour containing terms such as "explain." As Day has suggested, in commenting on Hilgard's objections to the apparent untestability of many Skinnerian explanations,

"For Skinner, his explanations are after all no more than instances of his own verbal behavior. If what he has to say, either by way of what is likely to be called an explanation or not, poses a problem to psychologists, then he sees only one avenue of interesting and effective comment as open: an analysis must be made of the variables that may be considered to have controlled his emission of the troublesome verbal behavior in question,"

and he goes on to claim that

... to fail to view the problem of explanation in this fashion, as inescapably an empirical and behavioral problem, is perhaps to miss the force of what may well be Skinner's major contribution to psychological thought." ([131], pp. 504-5).

That there *is* an interesting empirical-behavioral problem of this kind about explanation seems to me undeniable; that it is the *only* problem is a claim I would strongly oppose. But in any case, Skinner himself does not adopt this approach in stating his objections to mentalistic explanation, and neither shall I in criticizing those objections.<sup>6</sup>

## II

(1) *Mentalism lacks explanatory power.* This objection contains several distinct, though related elements, and I shall discuss each one separately.

(i) There is the complaint that mentalistic explanations are in some way “incomplete,” in that to explain behavior in terms of mental items is only to immediately raise the question of what it is that explains the mental item itself.<sup>7</sup> Thus, in “Behaviorism at Fifty,” one of Skinner’s “methodological objections” to mentalism is stated as follows:

“We may object, first, to the predilection for unfinished causal sequences. A disturbance in behavior is not explained by relating it to felt anxiety until the anxiety has itself been explained. An action is not explained by attributing it to expectations until the expectations have in turn been accounted for.” ([381, p. 240).

And the general spirit of this objection is well-expressed in his use of the phrase “mental way stations,” to suggest the unsatisfactory nature of the explanatory variables proffered by the mentalist.

As such, this is hardly a powerful objection, since it is in general true that we can always ask a further explanatory question concerning the variables provided in answer to the initial request for explanation. It is hardly a defect peculiar to mental items that we can always go on to investigate what it is that explains them. (See Hospers [211 on this general point).

A somewhat stronger form of (1)(i), which may perhaps be what Skinner sometimes has in mind, would be this: mentalistic explanations are always “incomplete,” because mental items can always be explained by reference to preceding stimuli and histories of reinforcement, whilst the converse does not hold. The trouble with this is that, first, it is by no means clearly true, and second, that even if it were, it would not show that mentalistic explanations have no explanatory power. The truth of the claim would, for example, be undermined if some mental items were “epiphenomena” of neurophysiological states, which were themselves not wholly explicable by stimuli and histories of reinforcement - a possibility that can hardly be ruled out, given the role of genetic factors. And even if the claim were true, it would hardly follow that mentalistic explanations had no explanatory power, unless one were prepared also to argue, say, that chemical theories have no explanatory power, on the grounds that chemical phenomena can be explained by physics, whilst the converse does not hold.<sup>8</sup>

(ii) The “incompleteness” arguments involved in (i) assume, in effect, a causal model in which mental items occur as “intervening” variables, in a sequence beginning with an environmental variable, and ending with a behavioral response: thus, in *Science and Human Behavior*, Skinner discusses mentalistic explanations in terms of

“... a causal chain consisting of three links: (1) an operation performed upon the organism from without — for example, water deprivation; (2) an inner condition - for example, physiological or psychic thirst; and (3) a kind of behavior - for example, drinking.” ([331, p. 34).

But in more recent works, a somewhat different model seems to be assumed, and an argument is given against mentalism’s explanatory power which would make the “incompleteness” argument redundant. For now the claim is that mental items are simply *by-products* of the same variables that are responsible for the behavioural responses. In commenting on an argument by Blanshard ([61, p. 331) that it would follow from the behavioristic position that Hitler’s hatred of the Jews contributed nothing towards his orders to have them exterminated, Skinner suggests that there is nothing absurd about this consequence, and recommends the view

“that both public persecution of the Jews and the private emotional responses were the result of Hitler’s environmental history.” ([381, p. 257).

And in *Beyond Freedom and Dignity*, the “by-product” claim is frequently made, as in the following passage:

“What is felt when a person protests is usually called resentment, significantly defined as ‘the expression of indignant displeasure’, but we do not protest *because* we feel resentful. We both protest *and* feel resentful because we have been deprived of the chance to be admired or receive credit.” ([39], p. 54: see also pp. 37, 72, 93, and 110).

If sustained, the claim would presumably establish mentalism’s lack of explanatory power; but Skinner produces neither experimental evidence nor adequate arguments to support it. The only argument for it that I can find in his work seems a poor one, and curiously it occurs in *Science and Human Behavior*, where the “intervening link” model is generally adopted. Thus, to show that “thinking of dinner” is a *by-product* of a ringing bell, and not itself a *cause* of salivation, he says:

“We cannot demonstrate that thinking of dinner will lead to salivation regardless of any prior event, since a man will not think of dinner in the absence of such an event.” ([33], p. 279).

But it would be absurd to require that someone who claimed that thinking of dinner was causally related to salivation (and not a by-product, with salivation, of a ringing bell) should establish that thinking of dinner itself had no cause. For it is surely not true that for one event, A, to be a causal condition of another, B, A must itself have no cause.<sup>9</sup>

(iii) Perhaps more important to Skinner than any of the preceding arguments for (1) is the claim that mentalistic explanations are defective because they fail to provide us with the means of predicting and controlling human behavior; and this accords with his general tendency to promote a “technological” view of the science of human behavior. (See, e.g., [39], chapter 1).

The issues involved in this claim are complex, and I shall confine myself to a brief discussion of only two of them. First, there is the general, philosophical problem of the relationship between explanation and prediction/control; and second, the question of whether it is in fact true that, if our concern is the control or prediction of human behavior, mentalistic explanations are less useful than their Skinnerian alter natives.

Skinner’s attitude to the first of these issues is not entirely clear. Sometimes, he may be read as supporting the view that to be able to *explain* behavior is basically equivalent to being able to predict or control it. At other times, it seems that he may be adopting a somewhat different position, of roughly the following kind: amongst the various causal conditions that may be functionally related to behavior, the most *important* are those which can be utilized for its prediction or control.

I think that the voluminous literature on the philosophy of explanation has shown the former of these views to be incorrect<sup>10</sup> - and, as I shall point out later, some of Skinner’s recent remarks concerning physiological explanations are probably incompatible with it. The second view is more interesting, and perhaps more defensible. Clearly, it rests upon the adoption of some account of “importance,” and for Skinner the choice is obvious: since the objective of a science of human behavior is to solve the problems confronting the human species, and since this can only be done by changing human behavior, the important variables are those we can use to do this.<sup>11</sup>

Of course, this is not the only possible view about the objective of a science of human behavior,<sup>12</sup> and Skinner neither explicitly criticizes alternative views, nor does he always display a particularly tolerant attitude to those whose concerns may be somewhat different from his own. But leaving this question aside, I would like to cast some doubt on the claim that, if our concern is to predict and

control behavior, mentalistic explanations are either useless, or at least less useful than their Skinnerian counterparts. It often seems that Skinner regards this claim as *obviously* true, whereas I believe that plausible counter-examples can be stated, both for prediction and control. I shall consider here just one example, concerning control. The “behavioural problem” is to change the way in which a particular man behaves towards women, from a manner displaying attitudes of male superiority and beliefs in “psychological differences between the sexes,” to one based upon a total rejection of those attitudes and beliefs.

There seem to me two important features of this case. The first is that the required change involves an enormously (indeed I would say *indefinitely*) large number of specific, often relatively “fine-grained” behavioral responses, both verbal and non-verbal. A considerable part of the man’s vocabulary has to be adjusted - e.g. “going with a woman to dinner”, rather than “taking a woman out to dinner” - along with behavior required by the now-rejected social etiquette, and the way in which the man previously distinguished between men and women in the types of facial expression with which he greeted them, the sorts of jokes he found it appropriate to tell, and so on. (I think anyone who has experienced this type of change will realize how wide-ranging and systematic it is.)

The second feature is that there is a good *prima facie* case for regarding the behavioral responses involved as being at least partly a function of the man’s beliefs and feelings, even if these are regarded only as “intervening links” (see pp. 57-58 above) - e.g. the man believes it to be his task to order food in the restaurant, feels anxious when desired responses are not forthcoming, and so on.

Now I would suggest that if we are concerned to bring about the desired changes in the man’s behavior, it may well be easier to do so by attempting to alter the relevant mental items, rather than concentrating directly on shaping up the behavior. For if the mental items are changed, the whole range of “appropriate” behaviour in (indefinitely) varied circumstances may well emerge, whereas we may otherwise be faced with the staggering task of shaping each behavioral response-class separately. Thus, instead of attempting to set up suitable contingencies of reinforcement for the behavior-changes, we should engage in argument and other interaction with the man, encouraging him to become more conscious and self-critical of the way he acts and thinks, with the primary emphasis on changing his convictions, feelings, etc., rather than on the direct manipulation of the behavior.

No doubt any Skinnerian worthy of that name will be able to provide objections to this example, and I shall comment briefly here on two possible counter-moves that he might make. The first would be to fall back upon “response-generalization,” and claim that this would substantially reduce the number of behavioral responses that would require individual shaping. I am inclined to think that this is a question-begging move, since the crucial problem is the criterion of similarity between the relevant

responses, and in my example the similarities are not behavioral, but, as it were, conceptual: thus all that the examples I gave of required behavioral changes have in common is their expression of certain beliefs and attitudes, and not anything that can be defined in behavioral terms. Thus there is no particular reason to believe that the “generalization” will take the required form.

The second move would be to claim that, suitably re-described, my counter-example could easily be accommodated within a Skinnerian framework: after all, all that is really involved in the proposed program of change is the manipulation of certain environmental variables, in this case predominantly verbal, which affect the man’s behavior (see, e.g., Skinner [36], p. 239). To this I would reply in two stages. First, it is still true that the direct concern in the example is with changing the relevant mental items, and not moving immediately to the desired changes in behavior. Second, even if a suitable re-description of the example can be achieved, why prefer this to the way I have presented it? No doubt the reply would be that my description tends to “divert our attention from the main problem, namely changing *behavior*” - but I will leave this point until the later examination of it as objection (4).

(iv) The final strand in the objection to mentalism’s lack of explanatory power which I will comment on is as follows: Skinner occasionally argues that mentalistic explanations are deficient, on the grounds that the very behavior we wish to explain provides our only means of identifying the presence of the “explanatory” mental items. Thus, in *Contingencies of Reinforcement*, he comments that

“... it is not good scientific practice to explain behavior by appealing to independent variables which have been inferred from the behavior thus explained, although this is commonly done, particularly by psychoanalysts, cognitive theorists, and factor analysts.” ([38], p. 264).

and I take it that the following remark in *Beyond Freedom and Dignity* can be regarded as pointing to this “practice”:

“We say that a person behaves in a given way because he possesses a philosophy, but we infer the philosophy from the behavior and therefore cannot use it in any satisfactory way as an explanation, at least until it is in turn explained.” ([39], p. 30).

The objection may be stated either at the “practical” or the “philosophical” level: either as a problem about the usefulness of such explanations for prediction/control, or as a claim that the explanations do not meet the logical requirement of the “independent identifiability” of “cause” and “effect.” I doubt myself whether this is a logical requirement: but even if it is, I do not think that mentalistic

explanations systematically fail to meet it.<sup>13</sup> For although it may be true that our only evidence for (other people's) mental items is behavioral, there is no good reason to believe that, in explaining any *particular* piece of behavior, it is *that* behavior which provides our only evidence for the explanatory mental item. For example, I may attempt to explain someone's dropping a cup in terms of her feeling of anxiety, and use the way she was moving her hands prior to that event as evidence for the mental item; the evidence is behavioral, but not the same behavior as that which is being explained.

(2) *Mentalism involves the employment of "theories"*. Whereas in examining (1) above I was mainly concerned to show that the objection did not hold, in examining (2) I shall be considering why it is that this should be regarded by Skinner as a good objection. In doing so, I shall also discuss Skinner's attitude towards the natural sciences, and his support of the "naturalist" position, according to which it is proper to adopt the methodology of the natural sciences in the analysis of human behavior.<sup>14</sup> And here we are immediately faced by a somewhat paradoxical situation; for Skinner combines a vigorous defense of naturalism, (see, e.g., [33], chapter 2), with an equally vigorous refusal to adopt one of the characteristic features of the natural science, namely the employment of *theories* (see, e.g., [34]). Indeed, when this combination is applied specifically to the issue of mentalism, it might well seem that Skinner is committed to the view that mentalism is objectionable both because it employs typically "pre-" or "un-scientific" concepts of an animistic/vitalistic kind, and also because it makes a typically "scientific" use of theories.

For there is no doubt that mental items are "theoretical" in the sense of the term "theory" that Skinner explicitly adopts, and little doubt that this sense of "theory" is pretty much the same as that involved in the claim I have (implicitly) made above, viz, that the employment of theories is a characteristic feature of the natural sciences. Skinner's definition of "theory" was given by him in his paper "Are Theories of Learning Necessary?" ([34]), and in the Preface to *Contingencies of Reinforcement* he quoted and commented upon the definition as follows:

"The word 'theory' was to mean 'any explanation of an observed fact which appeals to events taking place somewhere else, at some other level of observation, described in different terms, and measured, if at all, in different dimensions' - events, for example, in the real nervous system, the conceptual system, or the mind." ([38], p. vii).

Cf. the following remarks of Carl Hempel:

"Scientific systematization is ultimately aimed at establishing explanatory and predictive order among the bewilderingly complex 'data' of our experience, the phenomena that can be 'directly observed' by us. It is a remarkable fact, therefore, that the greatest advances in

scientific systematization have not been accomplished by means of laws referring explicitly to *observables*, i.e., to things and events which are ascertainable by direct observation, but rather by means of laws that speak of various *hypothetical* or *theoretical*, *entities*, i.e., presumptive objects, events, and attributes which cannot be perceived or otherwise directly observed by us.” ([18], p. 177).

But before examining Skinner’s attitude towards theories (which include, of course, both mentalistic and neurophysiological kinds), I will make a few comments upon his naturalism.

In supporting this position, and employing it against mentalism, Skinner often appeals to “famous episodes” in the history of science that mark out the long struggle during which explanations of natural phenomena in terms of gods, spirits, souls, appetites, and vitalistic forces were eventually replaced by properly “scientific” ones. (See, e.g., [36], [37], and [39] chapter 1). The lesson we are meant to learn is that a science of human behavior will, similarly, only develop successfully when the employment of mentalistic explanations is abandoned. But the argument is unconvincing: for the characterization of scientific development implicit in it is largely misconceived, and the suggested analogy between the natural sciences and psychology precisely begs the question.

This characterization of the history of science, not unlike that espoused by 19<sup>th</sup> century French positivists, involves at least three assumptions. First, that there is a relatively clear distinction that can be made between “un-” or “pre-scientific” approaches, and a “scientific” one; second, that this distinction can be stated roughly in terms of the contrast between resorting to animistic or “metaphysical” explanations, and confining oneself to orderly arrangements of observable data which provide us with predictive/controlling power; and third, that there is a gradual, though uneven progress from the one approach to the other.

Against this, I would suggest that nearly everything we have learnt from the increasingly sophisticated work of historians of science counts strongly against these assumptions.<sup>15</sup> It is perhaps indicative of Skinner’s approach to these issues that, when he cites an example from Herbert Butterfield’s deservedly well-known *Origins of Modern Science*, he ignores the main arguments contained in the chapter from which the example is taken. Thus Skinner states that:

“Although physics soon stopped personifying things ... it continued for a long time to speak as if they had wills, impulses, feelings, purposes, and other fragmentary attributes of an indwelling agent. According to Butterfield, Aristotle argued that a falling body accelerated because it grew more jubilant as it found itself nearer home . . .” ([39], p. 8).

But he does not mention, amongst others, the following point made by Butterfield:

“It is astonishing to what a degree not only this theory [viz. Aristotle’s theory of motion - R.N.K.] but its rivals - even the ones which superseded it in the course of the scientific revolution - were based on the ordinary observation of the data available to common sense. And, as writers have clearly pointed out, it is not relevant for us to argue that if the Aristotelians had merely watched the more carefully they would have changed their theory of inertia for the modern one - changed over to the view that bodies tend to continue either at rest or in motion along a straight line until something intervenes to stop them or deflect their course. It was supremely difficult to escape from the Aristotelian doctrine by merely observing things more closely . . .” ([81, p. 4).

Yet this point is of far greater significance for Butterfield’s general characterization of the development of modern science than the “jubilation” example, though it is, of course, hardly compatible with the characterization implicit in Skinner’s “historical” argument for naturalism.<sup>16</sup>

In any case, this argument cannot be used against mentalistic explanations of human behavior without begging the question. For it does not follow from the fact that it is mistaken (or “unscientific”) to explain natural phenomena in animistic terms, that it is likewise mistaken or unscientific to explain human behavior in mentalistic terms: falling stones do not feel jubilant, but humans sometimes do. (For a similar point, see Chomsky [9], p. 19).

Returning now to the question of Skinner’s attitude to theories, and thus to mentalism as one variety of “theoretical” explanation, I shall proceed by making a few remarks on a passage contained in a fairly recently published paper, “The Inside Story.” This, I believe, contains some useful evidence concerning Skinner’s views that was not available to Scriven and Johnson in their interesting and important exchange on this issue (see [31] and [22])

“A behavioral analysis is essentially a statement of the facts to be explained by studying the nervous system. It tells the physiologist what to look for. The converse does not hold. We can predict and control behavior without knowing how our dependent and independent variables are connected. Physiological discoveries cannot disprove an experimental analysis or invalidate its technological advances.

“This is not to question the importance of physiology in a science of behavior. In a more advanced account of a behaving organism ‘historical’ variables will be replaced by ‘causal.’ When we can observe the momentary state of an organism, we shall be able to use it instead of the history responsible for it in predicting behavior. When we can generate or change a

state directly, we shall be able to use it to control behavior. Neither the science nor the technology of behavior will then vanish, however. Physiological manipulations will simply be added to the armamentarium of the behavioral scientist. Pharmacology has already foreshadowed this state of affairs. A drug changes an organism in such a way that it behaves differently. We may have been able to make the same change by manipulating standard environmental variables, but the drug now permits us to circumvent that manipulation. Other drugs may yield entirely new effects. They are used as environmental variables.” ([371], p. 283).

(i) Although he acknowledges the potential of physiological explanations for prediction and control, it is clear that Skinner does not regard their explanatory power as wholly dependent upon this. This suggests that Skinner does accept a distinction between (causal) explanation, and the ability to predict/control. (Cf. p. 59 above.) And it also suggests the legitimacy of a science of human behavior whose primary objective is not “technological.” (Cf. p. 59 above).

(ii) If (i) is correct, and I am right in thinking that at least some physiological explanations are, for Skinner, “theoretical,”<sup>17</sup> then it would seem that the objection that mentalistic explanations are “theoretical” cannot by itself have any weight.<sup>18</sup> What would be required would be independent arguments against their explanatory power, and I have already criticized Skinner’s attempts to provide these. Indeed, it is not clear to me that the passage quoted here is entirely consistent with a general opposition to “theory”: possibly, a change of view is indicated.<sup>19</sup>

(iii) It seems that Skinner is prepared to describe the use of drugs as changes in “environmental variables.” If this is so, it is no longer obvious what is so distinctive about the “environmental solution” to behavioral problems which Skinner advocates so strongly, e.g., in a paper with that title (see [38], pp. 50-71). For there, as elsewhere, the environmental solution is one in which the desired changes are brought about by changes in the contingencies of reinforcement - and giving people drugs is surely not an example of this approach. (One wonders whether one is aiming at *Walden Two* or *Brave New World*.)

Finally, it is worth noting that it is not altogether clear what precisely is the source of Skinner’s opposition to the employment of theories. Johnson argues that the issue is primarily an empirical one, concerning the most fruitful procedure to adopt in the “context of discovery.” Thus, in commenting upon the title of Skinner’s paper, “Are Theories of Learning Necessary?” ([34]), she states that

“The question should be asked, ‘Are theories of learning necessary to what?’ And the answer is, I think, ‘Are they necessary, in the context of discovery, to acquire scientific knowledge?’

“

And she suggests that

“Skinner believes that his method of scientific procedure, his method of discovery, is more fruitful than that of constructing a hypothesis, deducing theorems, submitting them to experimental check, or starting out with a preconceived model of behavior” ([22], p. 275).

There is no doubt that Skinner *is* opposed to the use of this “hypothetico-deductive method,” but this is surely a separate question from that of the employment of theories, i.e., a certain mode of explanation.<sup>20</sup> Unfortunately, Skinner himself fails to separate the questions (see the Preface to [38]). But there is no reason why the “hypotheses” of the hypothetico-deductive method should involve “theories,” in Skinner’s sense of the term: in Popper’s advocacy of this method, for example, a hypothesis is simply a (strictly) universal statement (see [29], chapters I-V).

I would suggest, instead, that Skinner opposes the use of theories because they involve a kind of explanation which is not useful for the purposes of prediction and control. Insofar as this objection applies to mentalistic explanations, I have already tried to show that the objection may not hold, and it seems from the passage quoted from “The Inside Story” that Skinner does not now regard it as holding of all physiological explanations. In any case, there seems no good reason why we should not be interested in explanation “for its own sake,” and there may well be something to be said for what seems to have been Francis Bacon’s view, namely that in the long run, a successful technology can only be sustained by the discovery of causal explanations which involve the description of “inner” processes, structures, and so on.

(3) *Mentalism tends to invoke “homunculi” or “inner agents,” regarding man as an “autonomous agent.”* In the first three chapters of *Beyond Freedom and Dignity*, as elsewhere, Skinner tends to associate mentalism both with an (“unscientific”) belief in human autonomy or freedom (a belief which is often expressed in talk of human “agency”), and with an invocation of “homunculi” or “inner agents” to explain behavior. My main purpose here will be simply to point out that these issues are quite distinct, and to defend mentalism against “guilt by association.” It seems to me perfectly proper to accept the use of both mentalistic explanations, and the concept of human agency, without thereby being committed to either a defense of human autonomy, or a belief in “inner agents.”

There are many different arguments against human autonomy, but the most important one in this context, and that which Skinner himself seems to adopt (see, e.g., [33], p. 116), is that free-will is incompatible with the truth of determinism and that determinism is (probably) true. By “determinism”

I mean here simply the claim that there is a set of (causally) sufficient conditions for any event. Of course, many philosophers have argued that freedom and determinism are not incompatible, and Skinner does not address himself to this question;<sup>21</sup> but even if we accept the argument against freedom, it does not follow that the use of mentalistic explanations is inconsistent with a denial of human autonomy. For the argument against freedom from the truth of determinism does not rest upon any specific characterization of the nature of the causal conditions involved - they can be mental as well as physical. (It seems to me that much of the polemic in *Beyond Freedom and Dignity* is misdirected: what is most valuable about the book is its attempt to show how wide-ranging is the change in our beliefs and concepts that is required if we reject the belief in human autonomy, and this point need not be combined with an attack on mentalism.)

Similarly, we can make legitimate use of a concept of human *agency* without this committing us to a belief in autonomy. As Goldman has suggested ([161, pp. 78-85), we talk naturally in terms of agency when we regard causal relations as obtaining between *things*, rather than *events*. He himself presents an analysis of human agency in terms of behavior which is explained at least partly by reference to the beliefs and desires of the person concerned; but we might instead adopt a more general concept of agency which is applicable to both living and non-living objects. E.g., we might suggest that something be regarded as an “agent” if, with respect to some phenomenon that requires explanation, it is changes in that thing that are causally responsible for it.

As for the invocation of “homunculi,” one need only note that there is a considerable difference between regarding humans as agents, and explaining their behaviour in terms of an “inner agent,” a “little man” within us. Skinner’s objections to this practice seem well-founded (see, e.g., [36] and [37], *passim*) but neither the use of mentalistic explanations, nor the talk of human agency, commits us to it.

(4) *Mentalism distracts our attention from behavior.* A typical example of this objection is contained in the following passage:

“The dimensions of the world of mind and the transition from one world to another do raise embarrassing problems, but it is usually possible to ignore them, and this may be good strategy, for the important objection to mentalism is of a very different sort. The world of the mind steals the show. Behavior is not recognized as a subject in its own right. In psychotherapy for example, the disturbing things a person does or says are almost always regarded merely as symptoms, and compared with the fascinating dramas which are staged in the depths of the mind, behavior itself seems superficial indeed. In linguistics and literary criticism what a man says is almost always treated as the expression of ideas or feelings. In

political science, theology, and economics, behaviour is usually regarded as the material from which one infers attitudes, intentions, needs, and so on. For more than twenty-five hundred years close attention has been paid to mental life, but only recently has any effort been made to study human behavior as something more than a mere by-product.” ([39], p. 12).

Presumably, Skinner’s main reason (at least now), for objecting to approaches which divert our attention from behavior is that the human species is in an undesirable condition, and that any successful attempt to cure its problems must involve systematic changes in human behavior. But it does not follow from this that anyone who employs mentalistic explanations of behavior is thereby diverted from an interest in human behavior itself, and the possibilities of changing it. Of course, Skinner’s point in the quoted passage, and elsewhere, is not so much that mentalism *necessarily* has this consequence, but that, as a matter of fact, it usually does. But even if this is so, the remedy is surely not to abandon mentalism, unless there are other, independent reasons for doing so, and so far I have argued that the reasons given by Skinner are by no means conclusive.

In any case, I rather doubt whether mentalism has had, or does have this “undesirable” effect. It is not quite clear how one would establish this point, either way, but certainly the somewhat sweeping remarks made by Skinner in the above passage seem rather hard to justify. It is true, e.g., that many economists and political scientists use behavior in order to infer “attitudes, intentions, needs, and so on,” but it is surely not true that this is the *only*, or even the main interest they have in behavior. Similarly, though many psychotherapists do regard “the disturbing things a person does” as “symptoms,” to claim that they regard them *merely* as symptoms seems less plausible: after all, to regard something as a symptom does not mean that one is uninterested in removing it. Of course, to describe the “disturbing things” as “symptoms” may well imply the belief that certain ways of removing them may not be very effective, in the long run; but this surely suggests a strong, rather than a weak, concern with behavior.

It is also worth pointing out, here, that the fact that the solution to social problems *involves* changes in human behavior does not entail that one should *only* be concerned to study ways of bringing about such changes. One might argue, e.g., that the ultimate concern is with the amount of human happiness, and that this is at least partly a matter of generating circumstances in which people have certain pleasurable feelings and avoid unpleasant ones: thus, when people do not have enough to eat, they are not only not eating, but feeling pain. (And it seems to me unlikely that any analysis, such as Skinner’s, which contains no attempt to examine the political and economic structures which are responsible for poverty and hunger, will be successful in providing a solution to those problems).

### III

I hope by now to have shown that the arguments against mentalism advanced by Skinner in the objections so far discussed (i.e. (1) - (4)) are, at best, inconclusive. It might seem that this is not an entirely disastrous state of affairs, since there remains objection (5), that *mentalism involves a dualistic ontology of the "mental" and the "physical"*. However, quite apart from the question of how powerful an objection this is, there is the fact that Skinner often denies that he attaches much significance to it. One instance of this denial has already been cited (see the first sentence of the passage quoted on the preceding page), and another occurs in the following comment:

“Nor can we escape from the primitive features [of ‘homunculism’ - R.K.] by breaking the little man into pieces and dealing with his wishes, cognitions, motives, and so on, bit by bit. The objection is not that these things are mental but that they offer no real explanation and stand in the way of a more effective analysis.” ([36], p. 222: see also p. 221).

Both these passages are consistent with what seems to me a general tendency on Skinner’s part to avoid conducting a defence of radical behaviorism on what might be called “traditional metaphysical” grounds. The only passage I have been able to discover where this tendency is not displayed is the following one, from *Science and Human Behavior*:

“Modern science has attempted to put forth an ordered and integrated conception of nature. Some of its most distinguished men have concerned themselves with the broad implications of science with respect to the structure of the universe. The picture which emerges is almost always dualistic . . . Such a point of view . . . obviously stands in the way of a unified account of nature. The contribution which a science of behavior can make in suggesting an alternative point of view is perhaps one of its most important achievements.” ([33] p. 258).

Significantly, this passage comes near the beginning of the chapter in which Skinner discusses “Private Events in a Natural Science;” clearly he feels that one of the advantages of private events, as opposed to mental ones, is their physical status, and thus their being able to be incorporated in a “unified account of nature,” avoiding dualism.

Before concluding by commenting briefly on some of the problems in Skinner’s treatment of private events, it is worth noting that it is perhaps fortunate that private events seem to have this advantage, since several of the objections to mentalism must surely apply to private events also. Thus, explanations involving reference to private events would also be “incomplete,” and unsuitable for use

in the control of behavior, and there would be no less reason for regarding private events as “by-products” than there would be for their mentalistic counterparts. Our evidence for them would also consist in the behavior we are trying to explain, and a concern for private events might equally divert our attention from behavior.

This is, of course, hardly surprising since, as I noted at the beginning, many items that would traditionally be termed “mental” turn up again in Skinner’s account as private events; the only difference is, it would seem, that they are to be regarded as physical, and not mental. But even if there are very strong reasons for rejecting a dualistic ontology (and thus for maintaining (5) as the main objection against mentalism), I do not think that Skinner has dealt sufficiently carefully with the problems involved in ascribing a physical status to items such as toothaches, pains, and various feelings.<sup>22</sup> That this is so is indicated by his tendency to characterize private events simply as those which occur “inside the skin” of the organism, and to suggest that the development of physiological techniques and instruments will enable us to observe them directly. Thus, in a note on “Private Stimuli,” Skinner claims that “it is . . . absurd to argue that because they occur inside the skin they have non-physical dimensions,” ([381], p. 242), and at the end of the chapter on “Private Events” in *Science and Human Behavior*, he expresses the hope that “the problem of privacy may, therefore, be solved by technical advances” ([331], p. 282).

But someone who rejects the view that, e.g., a toothache is physical does not do so merely on the grounds that this is a stimulus which occurs “inside the skin.” The inflammation of the nerve that is responsible for the toothache also occurs inside the skin, as does the circulation of the blood, and the beating of the heart. There is nothing strange about claiming that these last three events may be observed by the use of suitable instruments, but it remains philosophically problematic whether it makes much sense to talk of this in the case of the toothache, or to ascribe a physical status to it. (See, e.g., the discussion of Malcolm’s contribution to [40], pp. 155-162, and Begelman [4]).

Indeed, Skinner has recently claimed that the evidence from

“recently developed physiological techniques . . . simply points to the fact that what is experienced introspectively *is* a physical condition of the body, as a behavioristic theory of knowledge has always contended.” ([38], p. 262: my italics).

The evidence is not cited: but, as many contemporary materialists have argued, in defense of the identity-thesis (see, e.g., [2], chapter 6), the italicized “is” in this passage requires philosophical justification, and cannot be established by empirical evidence alone. And this last point, when

generalized to apply to many of Skinner's claims, is an appropriate one with which to conclude.

---

### Notes

1 This is a considerably revised version of a contribution to the symposium on "Private Events in the Control of Behavior" at the American Psychological Association's convention in Washington, D.C., September 1971. I am particularly grateful to Willard F. Day for his comments on an earlier draft.

2 For a lucid account of the distinctive features of "radical behaviorism," see Day [12]. The differences between this, and other forms of behaviorism have been noted and discussed in Scriven [31], in Malcolm's contribution to [40], and in MacCorquodale [27].

3 These difficulties, and the relationship between Skinner's and Wittgenstein's approaches to them, are discussed in Day [13] and Begelman [4].

4 Both Johnson ([22]), and Day ([14]), have argued that Scriven ([31]) is wrong to regard these translations as analyses of meaning, or as philosophical "reductions."

5 Another purpose is that of exegesis. Wherever possible, I have tried to give references to passages which illustrate any view which I attribute to Skinner, and some of these are quoted in the text. Where no such references are given, this is usually because I am relying on general impressions I have formed from reading Skinner's work: naturally, these may be erroneous. It may be objected that my practice of examining each element in Skinner's anti-mentalism separately leads to a failure to capture what might be termed the "general spirit" of his opposition, but I believe that the approach adopted here has sufficient advantages to outweigh this possible defect.

6 An implicit assumption of these last few paragraphs is, of course, that philosophy is an essentially evaluative discipline. Unfortunately, the term "evaluative" is often now employed to indicate that moral or ethical issues are at stake, but I am using it in its more general sense, as concerning questions of justification, of the (logical) adequacy of our grounds for believing or accepting or doing certain things. Thus Hilgard's "untestability" objection is an example of this "evaluative" approach: the claim is that Skinner's explanations fail to meet a certain standard that must be met by any rationally acceptable scientific explanation. But even in this more general sense of the term, some philosophers might still disagree that their discipline is evaluative. (For other examples of Skinner's treatment of "evaluative" questions, see the chapter on "Values" in [39], and the comments on logical argument in [35], chapter 18.)

7 I use the phrase "mental items" to refer indifferently to mental states, events, processes, and so on. Usually, the context will make it clear which, if any, of these is more appropriate than the others.

8 That these issues are by no means peculiar to psychology and behaviorism is well-illustrated by, e.g., Brian Barry's discussion of the claim that "values" have no explanatory power with respect to social phenomena. made on the grounds that "values" can always themselves be explained by "interests". See [3], pp. 96-98.

9 It is not quite clear how one should understand the phrase "regardless of any prior event": it might mean "no matter what event occurred earlier." But again, we do not have to show that A would be followed by B no matter what preceded A, in order to claim that A is a causal condition of B.

10 See, e.g., Scheffler [30], pp. 19-57, and the detailed discussion of the objections to the "logical symmetry" of explanation and prediction in the title essay of Hempel [18].

11 The general issue of criteria for selecting certain causal conditions as more "important" than others has recently been much discussed by philosophers of history: see, e.g., Dray [15], chapter 4, and White [41], chapter 4. For a somewhat different approach, see Blalock [5].

12 For an interesting discussion of the way in which this question affects the "dispute" between Skinner and Chomsky, see Kaufman [23].

13 The issues here are extremely complex, and there is, as yet, little agreement about them by philosophers of mind. See, e.g., Davidson [11], and Pears [28]. For a discussion of how the naturalism versus anti-naturalism debate may be affected by recent developments in the philosophy of science, see Keat [24].

15 Notable examples of this work include Crombie [10], Koyré [25], and Kuhn [26].

16 In fact, Butterfield does not produce textual support for the attribution of this view to Aristotle, and I am strongly inclined to think that it is mistaken.

17 Obviously, there is a problem here about what counts as “physiological.” But it seems to me that the kinds of “physiological” explanations discussed by Skinner in this passage must include at least some which make reference to events in the “real nervous system,” and these are clearly “theoretical” in the light of the passage quoted earlier from the Preface to *Contingencies of Reinforcement* ([38], p. vii, on p. 62 above). For a different view, see Johnson [22], p. 276.

18 Cf. Bloor [7], for an interesting comparison of arguments against mentalistic explanations with those used in the 19th century against “theoretical” explanations which involved reference to atoms.

19 Certainly, a more “hopeful” attitude is displayed towards the usefulness of physiological explanations in the passage quoted here from “The Inside Story” than in earlier comments: e.g. in [33], pp. 28-9.

20 For an interesting discussion of the distinction between the contexts of “discovery” and “justification,” and of the different ways in which scientists “arrive” at theories, see Achinstein [1], chapters 6 and 7.

21 A powerful attack on the “compatibility” view, together with a general discussion of the free-will problem, is contained in Honderich [20], chapter 5.

22 In any case, the concept of the “physical” is itself by no means unproblematic, as is shown by recent discussion of the criteria for “physical existence.” See, e.g., Harré [17] and Hesse [19].

## References

- [1] Achinstein, P., *Law and Explanation*, Oxford: Oxford University Press, 1971.
- [2] Armstrong, D.M., *A Materialist Theory of Mind*, London: Routledge and Kegan Paul, (New York: Humanities Press), 1968.
- [3] Barry, B.M., *Sociologists, Economists and Democracy*, London: Collier-Macmillan, 1970.
- [4] Begelman, D.A., “Wittgenstein and Skinner on Private Languages, Rules and Stimuli”, paper contributed to the symposium on “Private Events in the Control of Behavior” at the American Psychological Association convention in Washington, September 1971.
- [5] Blalock, H. M. Jr., “Evaluating the Relative Importance of Variables”, *American Sociological Review*, Vol. 26, No. 6, 1961.
- [6] Blanshard, B. and Skinner, B.F., “The Problem of Consciousness - a Debate”, *Philosophy and Phenomenological Research*, Vol. 27, 1967.
- [7] Bloor, D., “Is the Official Theory of Mind Absurd?”, *British Journal for the Philosophy of Science*, Vol. 20, 1970.
- [8] Butterfield, H., *The Origins of Modern Science*, London: G. Bell and Sons Ltd., 1957.
- [9] Chomsky, N., “The Case Against B.F. Skinner”, *New York Review of Books*, Vol. XVII, No.11, 1971.
- [10] Crombie, A.C., *Augustine to Galileo*, 2 vols., 2nd edn., London: Heinemann, 1959.
- [11] Davidson, D., “Actions, Reasons and Causes”, *The Journal of Philosophy*, Vol. LX, No. 23, 1963.
- [12] Day, W.F., “Radical Behaviorism in Reconciliation with Phenomenology” *Journal of the Experimental Analysis of Behavior*, Vol. 12, No. 2, 1969.
- [13] Day W.F., “On Certain Similarities between the *Philosophical Investigations* of Ludwig Wittgenstein and the Operationism of B. F. Skinner”, *Journal of the Experimental Analysis of Behavior*, Vol. 12, No. 3, 1969.
- [14] Day, W.F., “Beyond Bondage and Regimentation: a review of *Beyond Freedom and Dignity* by B.F. Skinner”, *Contemporary Psychology*, 1972, Vol. 17, pp. 465-469.
- [15] Dray, W.H., *Philosophy of History*, Englewood Cliffs, NJ: Prentice-Hall, 1964.
- [16] Goldman, A. I., *A Theory of Human Action*, New York: Prentice-Hall, 1970.
- [17] Harré, H.R., “The Structure of Existential Judgments,” *Philosophical Quarterly*, Vol. 15, 1965.
- [18] Hempel, C.G., *Aspects of Scientific Explanation*, New York: The Free Press, 1965.
- [19] Hesse, M., “Laws and Theories”, *The Encyclopedia of Philosophy* (ed. Edwards, P.), Vol. 4., New York:

Macmillan and The Free Press, 1967.

[20] Honderich, T., *Punishment: The Supposed Justifications*, London: Hutchinson, 1969, and Harmondsworth: Penguin Books, 1971.

[21] Hospers, J., "What is Explanation?", in *Essays in Conceptual Analysis* (ed. Flew, A.), London: Macmillan, 1956.

[22] Johnson, R.J., 'A Commentary on "Radical Behaviorism"', *Philosophy of Science*, Vol. 30, No. 3, 1963.

[23] Kaufman, A.S., "The Aims of Scientific Activity", *The Monist*, Vol. 52, No. 3, 1968.

[24] Keat, R.N., "Positivism, Naturalism and Anti-Naturalism in the Social Sciences", *Journal for the Theory of Social Behavior*, Vol. 1, No. 1, 1971.

[25] Koyré, A., *From the Closed World to the Infinite Universe*, New York: Harper, 1958.

[26] Kuhn, T., *The Structure of Scientific Revolutions*, 2nd edition, Chicago: University of Chicago Press, 1970.

[27] MacCorquodale, K., "On Chomsky's Review of Skinner's Verbal Behavior," *Journal of the Experimental Analysis of Behavior*, Vol. 13, No. 1, 1970.

[28] Pears, D.F., "Desires as Causes of Actions", in Royal Institute of Philosophy Lectures, Vol.1, *The Human Agent*, London: Macmillan, New York: St. Martin's Press.

[29] Popper, K.R., *The Logic of Scientific Discovery*, London: Hutchinson, 1959.

[30] Scheffler, I., *The Anatomy of Inquiry*, London: Routledge and Kegan Paul, 1964.

[31] Scriven, M., "A Study of Radical Behaviorism", in *Minnesota Studies in the Philosophy of Science*, Vol. 1. (Edited by Feigl, H. and Scriven, M.), Minneapolis: University of Minnesota Press, 1956.

[32] Skinner, B.F., "The Operational Analysis of Psychological Terms", *Psychological Review*, vol. 52, No. 5, 1945.

[33] Skinner, B.F., *Science and Human Behavior*, New York: The Free Press, 1953.

[34] Skinner, B.F., "Are Theories of Learning Necessary?", *Psychological Review*, vol. 57, No.4, 1956.

[35] Skinner, B.F., *Verbal Behavior*, New York: Appleton-Century-Crofts, 1957.

[36] Skinner, "Behaviorism at Fifty", *Science*, 140, 1963, reprinted in [40] and [38]: page references to [38].

[37] Skinner, B.F., "The Inside Story", *Psychology Today*, April 1969, reprinted in [38]: page references to [38].

[38] Skinner, B.F., *Contingencies of Reinforcement: A Theoretical Analysis*, New York: Appleton Century-Crofts, 1969.

[39] Skinner, B.F., *Beyond Freedom and Dignity*, New York: Alfred A. Knopf, 1971.

[40] Wann, T.W. (Ed.), *Behaviorism and Phenomenology*, Chicago and London: University of Chicago Press, 1964.

[41] White, M., *Foundations of Historical Knowledge*, New York and London: Harper and Row, 1965.