

Individualism and Theories of Action

Steven Davis
Philosophy Department
Simon Fraser University

Abstract

In a recent series of articles Tyler Burge has presented arguments which cut against individualist theories of intentional states. In this paper I shall try to show what consequences Burge's arguments have for individualist theories of behavior. I shall take Jerry Fodor, who is one of the leading exponents of individualism in psychology, as representative of this view. First, I shall lay out one of Burge's arguments against individualist theories of intentional states; second, I shall describe the leading principles of Fodor's individualist metatheory for psychology; and lastly, I shall draw some of the consequences that Burge's arguments have for Fodor's theory of behavior.

* * *

Suppose that I believe that the sun sets over Vancouver Island and suppose that my report of the belief is *de dicto*.¹ It seems that I can have the belief, even if it were false and even if there were no sun nor Vancouver Island. That I have a belief and what belief that I have does not seem to depend upon the existence of anything except me. It appears that for me to have the beliefs that I do only requires that I have certain concepts or notions and the ability to combine them in certain ways. We might think that it follows from this that the conditions of individuation of beliefs depend on what is internal to an individual and not on the external objects to which an individual is related or on his social or linguistic community. One might further hold that a theory of beliefs, and intentional states in general, can be adequate without making reference to anything external to an individual. Borrowing a term from Tyler Burge, we shall call the views about the individuation and explanation of intentional states 'individualist' and theories which presuppose this view 'individualist theories of intentional states'(1979, p. 73).²

There are many things which I do which can be described in such a way that what I do does not seem to depend on anything except my existence, for example, my rubbing my thigh or my performing an action with the intention to relieve the pain in my thigh. We can extend Burge's notion to apply, as well, to the individuation of behavior and to theories of behavior. We can say that a criterion of individuation or a theory of behavior is individualist, if it is committed to the view that an adequate criterion or theory can be given without presupposing the existence of anything external to the individual. There is an obvious connection between individualist criterion of individuation and individualist explanatory theories; the latter presupposes the former. Hence, if it can be shown that the individuation of intentional states or of behavior is not individualist, then it follows that the explanatory theories of intentional states and of behavior cannot be individualist.

1. In what follows all the reports of intentional states are *de dicto*.

2. There are some states which are regarded to be intentional, such as knowing, which seem to depend for their existence on objects external to the individual. These states do not fall within the domain of individualist theories (Fodor, 1981, p. 228).

In a recent series of articles Tyler Burge has presented arguments which cut against individualist theories of intentional states. In this paper I shall try to show what consequences Burge's arguments have for individualist theories of behavior. I shall take Jerry Fodor, who is one of the leading exponents of individualism in psychology, as representative of this view. First, I shall lay out one of Burge's arguments against individualist theories of intentional states; second, I shall describe the leading principles of Fodor's individualist metatheory for psychology; and lastly, I shall draw some of the consequences that Burge's arguments have for Fodor's theory of behavior.

Let us begin then with Burge's argument against individualism. In this argument Burge presents a thought experiment in which he keeps constant the objects to which a subject is causally related and his internal states and changes only the linguistic practices of his surrounding community. In the actual situation a person, whom I shall call 'Oscar,' has the thought that he has arthritis in his thigh.³ That is, he uses 'arthritis' to refer to a rheumatoid condition in his joints and a similar pain which he has in his thigh. This use is contrary to the use of those in Oscar's speech community to whom he defers on matters about English, to his doctor's for example. Burge, then, describes a counterfactual situation in which Oscar's internal states remain the same. There is no change in the history of his stimulations, in his internal physical states, in his dispositions to behavior, when his behavior is described non-intentionally, and in the causal relations among them. The only difference is that those in his speech community to whom Oscar defers on linguistic matters use 'arthritis' in the way in which Oscar uses it mistakenly in the actual situation. In the counterfactual situation Oscar does not have the thought that he has arthritis in his thigh, for no one in the counterfactual situation has any notion of arthritis. They have a notion of a disease which can occur in the joints and in thighs which is not a notion of arthritis. Burge suggests that we could introduce the term 'tharthrititis' into English as it is actually spoken which would express the notion that 'arthritis' expresses in the counterfactual situation. We, then, could describe the thought that Oscar has in this situation, namely the thought that he has tharthrititis in his thigh. But the thought that Oscar has in the actual situation, the thought that he has arthritis in his thigh is not the same as the thought that he has in the counterfactual situation, namely the thought that he has tharthrititis in his thigh, for the two thought events do not have the same content, since arthritis is not the same as tharthrititis. So in the counterfactual situation Oscar lacks a thought he actually has and he has a thought that he actually lacks. The conclusion of this argument is that there are cases in which a person's internal states do not individuate his intentional states; reference must be made to the practices of the linguistic community of which he is a part. And the conclusion Burge draws from the thought experiments is that a criterion of individuation of intentional states cannot be adequate and be individualist. A necessary condition for their adequacy is that they make reference to objects and linguistic practices which are external to the subject.

Let us now turn to the consequences I think that Burge's thought experiment has for individualist theories of action. I shall take as representative of these views a theory of Jerry Fodor who is one of the leading exponents of individualism in psychology. Fodor holds that our ordinary ways of talking about and explaining actions, when made rigorous and systematic, form the basis of cognitive psychology. What I wish to show is that our ordinary views about actions do not presuppose an individualist criterion of the individuation of action and in so far as psychology adopts our ordinary views about actions its theories cannot be individualist. Before

3. Following Burge, I shall use 'actual situation' and 'counterfactual situation' where others might use 'actual world' and 'possible world'.

turning to my arguments for this, I shall lay out the principle doctrines of Fodor's views about how psychology should be done.

Fodor holds then that one way of giving an ordinary explanation of a particular human action is by appealing to a subject's intentional states which cause the action. And one of the goals of cognitive psychology is to utilize this form of explanation. Hence, Fodorian cognitive psychology takes as one of its goals the explanation of individual human actions and presupposes that some of these actions are caused by the subject's intentional states (1982, p. 100). These intentional states are representations which relate a subject to a content and a psychological theory which is committed to this is a representational theory of mind. In addition Fodor holds that representations, instantiated internally in a subject, are in a "language" of thought where only the formal properties, that is, the syntactic properties, of the "sentences" of the language play a causal role. Semantic properties of the representations, such as being true, having a referent or having a meaning, play no role in internal mental processes (1981, pp. 231). It is, then, the formal properties of a subject's internal representations which cause the subject's behavior and by appealing to these internal representations we can explain human behavior (1981, p. 239). On Fodor's view this is tantamount to the hypothesis that mental states and processes are computational.

Fodor maintains that the computational theory of mind entails a version of what Putnam has called "methodological solipsism" (1975, p. 136). On Fodor's view an adequate theory of human behavior which attributes intentional states to subjects need make no appeal to the subject's external environment, including the actual objects with which he is causally related or the social relations in which he is embedded. What explains the subjects behavior are the internal causal relations among formal properties of representational states (1981, p. 244). Fodor's commitment to methodological solipsism is a commitment to what I have called 'individualist' theories of intentional states and of behavior.

The identity conditions for representational states entailed by Fodor's methodological solipsism are formal: if a and b are representational state tokens of a subject who bears the same relation to a and b and a and b have the same syntactic properties, then $a=b$. Moreover, sameness of formal properties entails identity of causal powers, since the causal powers of representational states are contained in their formal properties.

Fodor, further, holds that particular representational states, including wants and beliefs, are token identical to internal physical states of the individuals which have them (1981, p. 9 and p. 145). In the case of humans intentional states are token identical to particular brain states. Consequently, if there is no change in the internal physical states of a subject, then there are no changes in his intentional states. It follows on Fodor's view that if there are no changes in a subject's internal physical states, then there is no change in his behavior, since his behavior is caused by his internal representational states.

The last principle is not one Fodor adopts. But there is nothing in Fodor's work which suggests that he would reject it. If a and b are rigid designators for bits of behavior and there is a cause of a which is possibly not a cause of b , then a is not identical to b . This principle is similar to part of Davidson's criterion for the individuation of event tokens (1969, p. 179). However, he would not accept the modal addition to the antecedent. Despite this, I shall call it 'Davidson's principle'. I take this principle to be plausible, since the cause of an event is what brings about an event. That is, the cause of an event is the origin of an event and the origin of an event is essential to it. This is parallel to the essential origin which Kripke claims for material objects (1980, p. 114). Kripke offers something like a proof for his claim about material objects which, I believe, can be applied to my claim about event tokens. And an event's causes being essential to an event entail what I have called Davidson's principle.

It follows immediately from Fodor's commitment to methodological solipsism that in an adequate psychological theory a subject's behavior cannot be described in such a way that it presupposes the existence of any object other than the subject. We cannot describe Oscar as taking aspirin to relieve the arthritic pain in his thigh, for that presupposes the existence of aspirin and arthritis, but we could describe Oscar as either performing an action which he intends to be a taking of aspirin which relieves the arthritis in his thigh or moving his body in a certain way. Neither seems to contravene methodological solipsism. The former does not, since correctly describing someone as having the intention to take aspirin to relieve arthritis in his thigh does not presuppose the existence of aspirin or arthritis. It might seem that the latter description is contrary to methodological solipsism, since it presupposes the existence of Oscar's thigh which on some views could be taken to be external to him. For the moment I shall suppose that Oscar's thigh is not external to him and that it is not ruled out by methodological solipsism. Methodological solipsism, then, limits psychological theories to descriptions of Oscar's movements of his body or to descriptions of his actions which are described intentionally.

Fodor suggests that his representational theory of mind accords with current theories in cognitive psychology (1981, p. 226). But limiting psychological theories to descriptions of actions which make reference only to a person's moving his body or to the intentional states which bring about the action is to propose a radical revision of current psychological theorizing. Moreover, I believe that Burge's thought experiment creates problems for Fodor's proposals for a cognitive theory of behavior even where the theory is limited in the descriptions it permits for behavior. I shall present two arguments against Fodor's theory which cut against the two sorts of descriptions which methodological solipsism seems to allow for actions. I shall begin with a problem for his theory where the descriptions of actions are descriptions of an agent's moving his body.

Let us suppose that Oscar rubs his thigh because of his beliefs that he has arthritis in his thigh and that rubbing a part of his body which suffers from arthritis will reduce the pain and his desire to lessen the pain in his thigh. We can say, then, that Oscar's action of rubbing his thigh is caused by his beliefs and his desire. In the counterfactual situation nothing changes about Oscar's internal states, that is, about his stimulation patterns, brain states, dispositions to behavior, and the causal relations among them. And on Fodor's theory since representational states are internal physical states, these, too, do not change. If we take Oscar's language of thought to be the language which he speaks, described non-semantically, then there is no difference between the actual and counterfactual situations in the sentences of his internal language. Hence, nothing should change about his behavior, since it is caused by his internal states. Oscar's rubbing his thigh in the actual situation should be identical to his rubbing his thigh in the counterfactual situation. However, in the counterfactual situation Oscar does not have the belief that he has arthritis in his thigh. He cannot have this belief, because he has no notion of arthritis. Hence, in this situation this belief cannot be a causal factor in his rubbing his thigh. But if we suppose that his rubbing his thigh is the same event token in the actual and counterfactual situation and in the counterfactual situation it is not caused by the belief that he has arthritis in his thigh, then this belief cannot play a causal role in the actual situation. Let us suppose that it does. Then, his rubbing his thigh in the actual situation has a cause which it fails to have in the counterfactual situation. But it follows from Davidson's principle that his rubbing his thigh in the actual and counterfactual situations cannot be the same act token. However, if we maintain that in the actual and the counterfactual situations we have the same act token, then we must give up its having a cause in one situation which it does not have in the other. Consequently, it seems that Fodor must give up the causal efficacy of beliefs and with it any hope of explaining actions by generalizing over intentional states. But this dooms the representational/computational theory

of mind, since the purpose of the theory is to allow for such generalizations, while providing an account of causal mental states and processes.

Fodor can hold onto the causal efficacy of beliefs and their explanatory role by denying that in the actual and counterfactual situations in Burge's thought experiments the subject performs the same actions. But it would follow that there can be changes in a subject's behavior, where this is not described by making reference to objects external to the subject, even though there is no change in the subject's internal states. It would not, then, be only the formal properties of the internal representational states which cause actions, since they do not change from the actual to the counterfactual situation. Hence, the representational/computational theory of mind which, Fodor argues, entails methodological solipsism must be abandoned, since beliefs which are not identical to any internal states can cause actions. It follows that to have a full account of these beliefs appeal must be made to the linguistic practices of the linguistic community of which Oscar is a part. Thus, reference must be made to objects which are external to the subject. As a consequence, on the assumption that behavior is caused by intentional states a theory which entails methodological solipsism cannot be an adequate theory of behavior.

In the example above Oscar's action was described as an action of his moving his body. We also obtain consequences which are unacceptable to individualist theories of behavior, if we consider an action of Oscar which is intentionally described. Let us suppose that Oscar performs an action which he intends to be an act of taking two aspirins to relieve the arthritic pain in his thigh. And let us suppose further that he performs this action, because he believes that he has arthritis in his thigh and that taking aspirin relieves the pain of arthritis and he desires to relieve the pain in his thigh. In this case, as in the example above, we can say that Oscar's action is caused by his beliefs and desires. In the actual and counterfactual situations in Burge's thought experiment there is no change in Oscar's internal states and hence, on Fodor's theory there should be no change in Oscar's behavior. But in the counterfactual situation Oscar does not have the belief that he has arthritis in his thigh, nor can he have the intention of performing an action which relieves an arthritic pain in his thigh. He can have no such belief or intention in the counterfactual situation, because he and everyone in his speech community lack the notion of arthritis. Hence, in the counterfactual situation Oscar does not perform an action which he intends to be a taking of two aspirin to relieve the arthritic pain in his thigh. Thus, Oscar does not perform the same action in the actual and the counterfactual situations, even though there is no change in Oscar's internal states. Since the same causes should have the same effects, it follows that either in the actual or the counterfactual situation the internal representational states which Fodor's theory attributes to Oscar cannot be the sole causal factor which brings about his actions and cannot, therefore, be used to give a complete explanation of his behavior. Once again appeal must be made to the linguistic community of which Oscar is a part to have an adequate account of his behavior.

I have considered two sorts of descriptions of Oscar's behavior: his moving his body and his performing an action with a certain intention. Perhaps, a way out of the difficulties that I have raised for Fodorian cognitive psychology is to describe Oscar's behavior as bodily movements, rather than either his moving his body or his performing an action with a certain intention. If we consider again the example in the previous paragraph we could take it that there is a bodily movement of Oscar, or a series of them, which is caused by his intentional states which are internal representational states of Oscar and the bodily movement or a series of them is actually contingently identical to his action which he performs with the intention that it be a taking of two aspirin to relieve arthritic pain. However, in the counterfactual situation we would have the same bodily movement, but not the same action, for Oscar, lacking any notion of arthritis, could not have an intention involving this notion and thus, could not perform an action with this intention.

It would, then, be the movements of Oscar's body, so described, which are caused by his intentional states. This does not mean, of course, that actions are not caused by intentional states. Rather, it has as a consequence that what action is caused by a particular set of intentional states can vary depending on changes in the non-internal environment of a subject. In the actual situation Oscar's beliefs and desire cause his performing an action which he does with the intention that it be the taking of two aspirin to relieve arthritic pain in his thigh, since this action is identical to movements of his body.

This will not do. In the actual and the counterfactual situations by hypothesis the representational states are the same, but the actions performed are not. In the actual situation the action performed is an action done with the intention that it be a taking of two aspirin to relieve arthritic pain. This action is not performed in the counterfactual situation. Since it is appeal to representational states which is supposed to explain actions and the representational states in the counterfactual and the actual situations are the same, there is no explanation for the difference in the actions between the actual and the counterfactual situations. Hence, the representational/computational theory of mind cannot be a complete theory of human action.

I have assumed that methodological solipsism allows descriptions of an agent's moving his body and movements of his body which make reference to parts of an agent's body. But I think that a variant of Burge's thought experiment can be used to show that methodological solipsism rules out such descriptions. Let us imagine that Oscar is as we described him in the actual situation, but that in the counterfactual situation he has no legs to rub and that his brain is hooked up to electrodes which produce in him the visual experiences about his legs which are identical to the visual experiences which he has in the actual situation. We suppose further that he has exactly the same beliefs and other intentional states and the same dispositions to behavior which he has in the actual situation. That is, his internal states are identical in the actual and the counterfactual situations. On Fodor's theory cognitive psychology has within its domain the formal properties of internal representational states which can be specified without making any reference to external objects. As Fodor puts it, "...[Representational states] have no access to the *semantic* properties of such representations, including the property of being true, of having referents, or, indeed the property of being representations *of the environment*." (1981, p. 231) But, then, a theory of such representational states cannot explain the difference between Oscar's actually rubbing his thigh and counterfactually believing falsely that he does, since there is no difference in his representational states. Hence, the theory cannot explain Oscar's rubbing his thigh and thus has no place for descriptions of actions which make reference to parts of Oscar's body.

Let me conclude by summarizing what I think that I have shown. On Fodor's view if psychology is successfully to provide a theory of behavior, then it must make reference to the beliefs, desires, and wants, that is, to the intentional states of a subject. Further, Fodor claims that this goal can only be achieved, if the working cognitive psychologist is committed to methodological solipsism. Burge's thought experiments can be taken to show that if Fodor is construing psychological theories to be more rigorous and precise versions of our ordinary want and belief explanations, then these theories must violate methodological solipsism by making reference to objects and linguistic practices external to an individual. My arguments show that if intentional states are the cause of our behavior, then theories of behavior cannot account for it fully and be individualist. That is, they contain descriptions of behavior which violate methodological solipsism. Moreover, and perhaps more importantly if the computational theory of mind entails methodological solipsism, as Fodor contends, then no computational theory can give us a complete account of behavior.

— Bibliography —

Burge, Tyler. (1979). Individualism and the Mental. Midwest Studies in Philosophy: Studies in Metaphysics, Vol IV, eds. Peter A. French, *et al.* Minneapolis: University of Minnesota Press, pp. 73-122.

_____. (1982a). Other Bodies. Thought and Object, ed. Andrew Woodfield. New York: Oxford University Press, pp. 97-120.

_____. (1982b). Two Thought Experiments Revisited. Notre Dame Journal of Formal Logic, XXII, 3, pp. 284-293.

_____. (1986a). Individualism and Psychology. Philosophical Review, XLV, 1, pp. 3-45.

_____. (1986b). Intellectual Norms and Foundations of Mind. Journal of Philosophy, LXXXIII, 2, pp. 697-720.

Davidson, Donald. (1969). The Individuation of Events. Essays on Actions and Events. Oxford: Clarendon Press, pp. 163-181.

Fodor, Jerry A. (1981). Representations. Cambridge, Mass.: MIT Press.

_____. (1982). Cognitive Science and the Twin-Earth Problem. Notre Dame Journal of Formal Logic, Vol.23, 2, pp. 98-118.

Kripke, Saul. (1975). Naming and Necessity. Cambridge, Mass.: Harvard University Press.

Putnam, Hilary. (1975). The Meaning of Meaning. Minnesota Studies in the Philosophy of Science, ed. Keith Gunderson. Minneapolis: University of Minnesota Press, pp. 131-193.