



Philosophical Review

Must Psychology Be Individualistic

Author(s): Frances Egan

Source: *The Philosophical Review*, Vol. 100, No. 2 (Apr., 1991), pp. 179-203

Published by: [Duke University Press](#) on behalf of [Philosophical Review](#)

Stable URL: <http://www.jstor.org/stable/2185299>

Accessed: 17/02/2011 14:53

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=duke>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Duke University Press and *Philosophical Review* are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Review*.

<http://www.jstor.org>

Must Psychology Be Individualistic?

Frances Egan

A widely held and intuitively plausible tenet of computational psychology holds that if two individuals are identical in all physical respects then they are psychologically identical as well. According to this doctrine, any differences between individuals not reflected in their internal physical states are not psychologically relevant, and should be ignored by psychological theory. This thesis—that psychological states *supervene* on physical states of the individual subject to which they are ascribed¹—has become known as *individualism*.²

Individualism has recently come under attack from a number of quarters. The most serious challenge is based on the semantic thesis that the meaning of names, natural kind terms, and perhaps all general terms, is at least partly determined by factors external to thought, in particular, by environment and social context. Largely due to Hilary Putnam's famous "Twin Earth" thought experiments and Tyler Burge's imaginative examples,³ "externalism" is now the received view in semantics.

Tyler Burge has argued that these thought experiments undermine the individualism thesis in psychology. Burge takes the thought experiments to show that the content of propositional attitudes (for example, beliefs, desires) does not supervene on non-intentionally specifiable properties or states of the individual subject. (For example, consider an English-speaking individual and his physically identical twin who uses the vocable "water" to refer

¹Roughly, states of type A are supervenient on states of type B if and only if there can be no difference among A states without a corresponding difference among B states.

²The term was first coined by Tyler Burge in "Individualism and the Mental," *Midwest Studies in Philosophy* 4 (1979), pp. 73–121.

³Hilary Putnam, "The Meaning of Meaning," reprinted in *Mind, Language, and Reality: Philosophical Papers, Vol. 2* (Cambridge, England: Cambridge University Press, 1975). Also Tyler Burge, "Individualism and the Mental," and "Other Bodies," in *Thought and Object*, ed. A. Woodfield (Oxford, England: Clarendon Press of Oxford University Press, 1982), pp. 97–120.

not to H₂O, but to the superficially similar but chemically different liquid which fills Twin Earth's lakes and rivers. The two subjects express beliefs with different propositional contents when they utter the form of words "Cold water is good to drink on a hot day.") In denying that psychology is individualistic, Burge is claiming that features of an individual's environment or social context—features not reflected in the individual's internal physical (that is, non-intentionally specifiable) states—are psychologically relevant.

In this paper I examine the evidence for and against individualism. I begin with an argument that purports to establish that *any* scientific psychology must be individualistic. I argue that this metaphysical argument fails to establish its conclusion. I then turn to Burge's arguments for the claim that individualism is false. I argue that the semantically based argument, and a second argument purportedly based on psychological practice, do not support Burge's contention that individualism in psychology is false. I conclude that, while there is no reason to think that psychology must be individualistic, consideration of the goals of psychological theorizing suggests that it most likely will be.

I. A METAPHYSICAL ARGUMENT FOR INDIVIDUALISM

Jerry Fodor has recently given an *a priori* argument for the claim that psychology should be individualistic.⁴ According to Fodor, "individuation in science is *always individualistic*."⁵ If psychology is to be scientific, he concludes, it should taxonomize mental states individualistically. Furthermore, he claims, "if mind/brain supervenience goes, the intelligibility of mental causation goes with it."⁶ I shall argue that Fodor's general claims about individuation in science are false—a non-individualistic psychology is not thereby "unscientific." Nor, I shall argue, must a psychology which violates mind/brain supervenience make a mystery of mental causation. In what follows I shall *not* argue that psychology is or should be non-

⁴*Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Mass.: The MIT Press, 1987), Chapter 2.

⁵*Ibid.*, p. 32.

⁶*Ibid.*, p. 42.

individualistic; I am concerned merely to show that Fodor's metaphysical argument does not establish that psychology should be individualistic.

Fodor's argument rests on an analogy. He defines two predicates, "is an H-particle at time *t*" and "is a T-particle at time *t*," such that a particle is either an H-particle or a T-particle at a particular time depending upon whether a particular dime of Fodor's is heads up or tails up at that time. By simply turning over his dime, Fodor can change every particle in the universe from an H-particle to a T-particle, and back again. He claims that a particle physics that counted such relational predicates among its explanatory resources would be "mad" because "particle physics, like every other branch of science, is in the business of causal explanation; and whether something is an H- (T-) particle is irrelevant to its causal powers."⁷

The argument depends on the claim that scientific theories individuate entities by their causal powers.⁸ For Fodor, *two things have the same causal powers if and only if they have the same causal consequences across nomologically possible contexts.*⁹ The problem with H- and T-particle theory, Fodor claims, is that it type-distinguishes entities which have the same causal powers. A genuinely scientific theory, he claims, would type-identify two objects which differ only in that one is an H-particle and the other is a T-particle.

Non-individualistic psychology—that is, psychology which violates mind/brain supervenience—is alleged to be analogous to H- and T-particle theory in respect of its taxonomic practices. According to Fodor, when the twins in the thought experiments utter the form of words "I desire a glass of water" their mental states have the same causal powers. Insofar as psychology is interested in causal explanation, he concludes, it should type-identify their mental states. And if, as Fodor is willing to concede to Burge, common-sense psychology type-distinguishes the two states,¹⁰ then

⁷Ibid., p. 33.

⁸Ibid., p. 34.

⁹Ibid., p. 44.

¹⁰At *ibid.* on p. 34, summarizing his argument, he says "... common-sense deploys a taxonomy that *does* distinguish between [the twins]." I think this concession by Fodor is rash (see argument next section).

the individuation criteria employed by scientific psychology and by its folk counterpart are just different.

Fodor considers the following objection: the twins' mental states have different causal powers, and hence are of different types, by virtue of producing different behaviors. Thus, for example, my desiring something to drink causes me to ask for a glass of water. My twin utters the same vocables, but does not ask for water—her thirst produces a request for a different liquid. Our bodily movements in producing these behaviors are the same, of course; but on the construal of behavior relevant for psychology the behaviors are different.¹¹

Fodor has two replies to this objection. The first argues as follows:

First reply: If this argument shows that my mental state differs from my twin's, it's hard to see that it doesn't show that our brain states differ too. My twin is in a brain state that eventuates in his uttering the form of words 'Bring water'. I am in a brain state that eventuates in my uttering the form of words 'Bring water'. If our uttering these forms of words counts as our behaving differently, then it looks as though our brain states differ in their behavioral consequences, hence in their causal powers, hence in the state types of which they are tokens.¹²

The intended force of this reply is unclear. Fodor seems to be claiming that the anti-individualist argument, if sound, would prove too much; in particular, it would show that the twins' brain states are type-distinct, which contradicts a presupposition of the thought experiment. He singles out as problematic the premise that the behavior of the twins is type-distinct, arguing that *if* this premise is granted, then it follows that their brain states are type-distinct also.

¹¹In taking the behaviors to be type-distinct, the proponent of this objection cannot base the distinction on the fact that the behaviors are caused by different intentions, since the argument would then be question-begging. Rather, the distinction must rest on the claim that, as a matter of fact, behaviors are at least sometimes individuated relationally, by reference to features of the context. Put another way, the type-distinctness of behavior is intended as the premise of an argument for the type-distinctness of mental states, rather than as a conclusion from it.

¹²*Psychosemantics*, p. 37.

Fodor's argument can be reconstructed as follows:

- (1) Suppose we count the behavior of the twins as type-distinct.
- (2) Behavior is a causal consequence of brain states.
- (3) So, the causal powers of the brain states of the twins are different (from 1 and 2).
- (4) Scientific theories individuate states/entities by their causal powers (general principle).
- (5) So, the brain states of the twins are type-distinct (from 3 and 4, together with the fact that neurology, the theory which individuates brain states, is a scientific theory).
- (6) But it is absurd to think that the brain states of the twins are type-distinct simply because they live in worlds with different liquids.¹³
- (7) Hence, we should *not* count the behavior of the twins as type-distinct.

A problem with this line of reasoning can be brought out by the following example. Suppose two telephone answering machines (M1 and M2) have been hard-wired to produce short computer-synthesized messages in response to incoming calls. M1, which has been designed for the singles market, responds, "I love you. Please leave a message." The other machine, M2, which was sold to the Isle of View Hotel, responds, "Isle of View. Please leave a message." The company that built the machines was quite pleased to get the order for M2 after having built a number of M1 models. The company's technical staff pointed out to the salesman that the same machine could produce both messages, since the two messages have the same underlying acoustical forms.

¹³By hypothesis the twins are molecularly identical. In keeping with the rest of the Twin Earth literature we are simply ignoring the fact that the brain is largely composed of water. The same point could be made using an example like gold—the point being that the twins do not differ in their brain states simply because they are causally related to different substances in their environments (or because they are members of different linguistic communities).

Now consider the following argument:

- (1a) Suppose we count the messages produced by the two machines as type-distinct.
- (2a) The messages are causal consequences of the physical states of the machines.
- (3a) So, the causal powers of the physical states of the two machines are different.

The argument could be continued, in the fashion of Fodor's argument, to the conclusion that the physical states of the machines should be type-distinguished, but since it would be absurd to do so, we should not type-distinguish the messages produced by the machines.

There is clearly something wrong with the above argument. The two machines are by hypothesis physically identical; they therefore have the same causal powers. But does it follow that the machines' "behavior" (the messages produced) must be type-identified? Surely not. Our best linguistic theory would assign different syntactic and semantic representations to the machines' behavior. A difference in their contexts of utterance, not a difference in their causal powers, accounts for the fact that the machines produce different messages. The type-individuation of these behaviors is sensitive to factors other than the causal powers of the thing that produces them. (If you are inclined to think that the messages are *ambiguous* between the two readings, the point still goes through. Syntactic theory assigns a type-distinct syntactic representation to each of the two readings; disambiguation, that is, choice between these two readings, is purely contextual.)

The answering machine argument purports to establish that type-distinguishing the messages is *inconsistent* with type-identifying the physical states (or causal powers) of the machines that produce the messages. But, no such inconsistency has been established: linguistic theory would type-distinguish the messages, and physical theory would type-identify the machines. The implicit assumption (let's call it A_1) that linguistic types supervene on the physical states of the device is simply false.

On the presumed construal of Fodor's first reply, his argument parallels the answering machine argument. It purports to establish that type-distinguishing the twins' behavior is inconsistent with

type-identifying their brain states. The argument fails to establish its conclusion without begging the question against the anti-individualist. The anti-individualist proposes to type-individuate behavior partly by reference to its context. In doing so he is *denying* that behavior supervenes on the causal powers of brain states; that is, he is denying that the twins must be in different brain states. Whether or not such a position is “scientific” (an issue to be addressed below), it is certainly not incoherent.

The inference from (1) and (2) to (3), in Fodor’s argument, is not valid without an assumption that behavioral types supervene on brain types, which is the very question at issue. (1) says that the twins’ behavior is *type*-distinct. The proper construal of (2) is that behavioral events (that is, *tokens*) are causal consequences of brain states. To conclude that the causal powers of brain states (that is, brain state *types*) are distinct is to assume that there can be no difference in behavioral type without a difference in brain state type, which is just to assume that behavioral types supervene on brain state types. (Call this assumption A_2 .) This assumption, of course, is exactly what the anti-individualist denies.

To summarize: type-distinguishing the behavior of the twins is claimed by Fodor to lead to the unacceptable conclusion that the twins are in different brain states. The conclusion of the *reductio* has been blocked: we need not type-distinguish the brain states of the twins on the assumption that their behavior is different because the inference from type-distinct behavioral effects to a difference in causal powers of the underlying brain states is not valid without the question-begging assumption that behavioral types supervene on brain types.

The failure of Fodor’s argument should give no comfort to the anti-individualist. The original objection claimed that the twins’ mental states are type-distinct because their behavior is type-distinct. This argument rests on the unwarranted assumption that behavioral types supervene on *mental* state types (A_3). Even if the behavior of the twins *is* type-distinct, it doesn’t follow that their mental states have different causal powers, nor that their mental states should be type-distinguished by psychology. The three arguments—the answering machine parody, Fodor’s reply, and the original objection—are formally analogous: each rests on an unwarranted supervenience assumption (A_1 , A_2 , and A_3 , respectively).

Fodor has a second reply to the objection that since the behavior of the twins is type-distinct their mental states should be type-distinguished. He extends the analogy between non-individualistic psychology and H- and T-particle theory, pointing out that an H- and T-particle theorist might justify his individuating practices by arguing that H- and T-particles do indeed have different causal powers: H-particles enter into (cause) H-particle interactions and T-particles enter into (cause) T-particle interactions. According to Fodor, this move is analogous to the claim that since the behavior of the twins is type-distinct, their mental states are type-distinct. Both moves, he claims, are illegitimate.

Fodor is right about this. Each argument assumes a problematic supervenience thesis. From the claim that H- and T-interactions are type-distinct events it doesn't follow that the particles posited as their causes must be type-distinct, unless we know independently that "interaction types" supervene on particle types (we might call this claim A_4). But let's consider what would be required to fix up the two arguments. In the twins case we need (1) an argument that the twins' behavior *is* type-distinct, and (2) an argument that the type-distinct behaviors are best explained by type-distinct *mental states* (rather than by, say, aspects of their different environments). Analogously, the H- and T-particle theorist needs (1) an argument that "H-interactions" and "T-interactions" are indeed type-distinct events, and (2) an argument that these type-distinct events are best explained by reference to type-distinct particles which cause them (rather than by reference to an aspect of their context—namely, the orientation of Fodor's dime). As I have been at pains to point out, both arguments lack (2), that is, an argument for the relevant supervenience thesis. What about (1), an argument that the effects are indeed type-distinct? Anti-individualists (notably Burge) have argued that folk psychology taxonomizes behavior in such a way that the behavior of the twins would count as type-distinct. The argument for this claim will be discussed below; what is important in the present context is that such individuating principles are alleged to be part of an established practice. By contrast, the move by the H- and T-particle theorist to type-distinguish "H-interactions" and "T-interactions" is completely *ad hoc*. In fact, the real problem with H- and T-particle theory is not that there is anything *in principle* wrong with the way it individuates entities (that is, "relationally"), but that the distinc-

tion between H- and T-particles seems to do no genuine explanatory work. Individuating entities as H- or T-particles would not increase the explanatory power of a conventional particle physics that failed to make such a distinction. H- and T-particle physics doesn't explain any new phenomena; it simply affords us a new description of the phenomena already explained by the conventional particle physics. If an analogous charge could be sustained against non-individualistic psychology, then it should indeed be rejected, but such a case could not be argued from *a priori* principles.

I would now like to examine the principle (that is, premise 4) at the heart of Fodor's argument for individualism—that a scientific theory must taxonomize entities or states by their causal powers. Fodor rests his claims about individuating practices in science on *two* general principles:

Methodological point: Categorization in science is characteristically taxonomy by causal powers. Identity of causal powers is identity of causal consequences across nomologically possible worlds. *Metaphysical point:* Causal powers supervene on local micro-structure.¹⁴

The methodological point, I shall argue, is compatible with non-individualistic taxonomic principles. Only the metaphysical point constrains individuation in science, but the metaphysical point is false.

The methodological point, by itself, doesn't specify the kinds of properties that can qualify as causal powers. The stipulation that identity of causal powers is identity of causal consequences across nomologically possible contexts provides only a criterion for *identity* of causal powers, and this only insofar as we know antecedently how causal consequences are to be individuated. The methodological point suggests that causal powers are nomic properties—they might plausibly be construed as the properties in virtue of which a state (event) is subsumed under causal laws.¹⁵ Under this construal the methodological point is compatible with the non-individualistic (contextual) individuation of mental states just in case non-super-

¹⁴*Psychosemantics*, p. 44.

¹⁵Fodor endorses this construal of causal properties in "Making Mind Matter More," *Philosophical Topics* 17 (1989), pp. 59–79.

venient mental properties can figure in causal laws. I argue below that there is no reason to think that they cannot.

Fodor's metaphysical point *does* put constraints on the kinds of properties that can count as causal powers. The following is a plausible interpretation of the metaphysical point: where T1 is a theory which specifies the micro-structure of entities individuated by T2, T2 recognizes no difference in causal powers unless there is a difference in causal powers recognized by T1. In other words, the causal powers postulated by the macro-level theory *supervene on* the causal powers postulated by the micro-theory. If Fodor's "metaphysical point" is correct, it would seem to rule out a difference in the causal powers of subjects' mental states without a difference in the causal powers of their brain states; that is, it would seem to rule out type-individuating mental states as the anti-individualist suggests.

The metaphysical point, however, is surely false. It must be false for any fundamental particle theory, since fundamental particles are presumed not to have an underlying micro-structure. The metaphysical point, I shall argue, is false for other domains of science as well.

The rationale for the metaphysical point seems to be that if a theory does not taxonomize by reference to local micro-structure its "laws" are not causal laws. Indeed, it will be recalled that Fodor says "if mind/brain supervenience goes, the intelligibility of mental causation goes with it."¹⁶ According to Fodor, for the mental states of the twins to have different causal powers there would have to be a *mechanism* connecting (the causal powers of) each subject's mental state with relevant features of the environment, without affecting the subject's physiology (since the two subjects are, by hypothesis, physiologically identical). But, Fodor claims, "there is no such mechanism; you can't affect the causal powers of a person's mental states without affecting his physiology."¹⁷ To avoid postulating "crazy causal mechanisms" (or action at a distance), he concludes, we must say that the twins' mental states have the same causal powers, and any scientific psychology should type-identify them.

¹⁶*Psychosemantics*, p. 42.

¹⁷*Ibid.*, p. 39.

The idea seems to be that if the twins' mental states are type-distinct (have different causal powers) in virtue of the fact that they live in worlds with different liquids, then the liquids (H₂O and XYZ respectively) must *cause* their different mental states. But the only way the environment can affect their minds is by affecting their bodies. Since the twins are, by hypothesis, physically identical, to claim that their mental states are type-distinct is to invoke "action at a distance."

Fodor's argument, however, simply misconstrues the anti-individualist's position. The proper construal of the anti-individualist claim is not that differences in the environments *cause* a difference in the twins' mental states, but rather that environmental (contextual) factors are relevant to the *type-individuation* of their mental states. Our theories, and the taxonomies employed by them, are not caused by the world in any direct sense. In claiming that psychological taxonomy is sensitive to contextual factors the anti-individualist is not postulating any mysterious causal processes.

Burge puts the relevant distinction nicely:

Variations in the environment that do not vary the impacts that causally 'affect' the subject's body may 'affect' the individuation of the information that the subject is receiving, of the intentional processes he or she is undergoing, or of the way the subject is acting. . . .

Information from and about the environment is transmitted only through proximal stimulations, but the information is individuated partly by reference to the nature of normal distal stimuli. Causation is local. Individuation may presuppose facts about the specific nature of a subject's environment.¹⁸

Fodor has, therefore, given no cogent reason why scientific theories cannot count contextual features as relevant to taxonomy. There is clear evidence from the sciences that they often do. Geology distinguishes between igneous and sedimentary rocks on the basis of their origins, not their micro-structures. Biology often type-distinguishes anatomical structures which, though identical in their micro-structures, have distinct biological functions or evolutionary histories. Moreover, some biological laws are defined over *species*—for example, Cope's Law, which says that unspecialized

¹⁸Tyler Burge, "Individualism and Psychology," *The Philosophical Review* 95 (1986), pp. 16–17.

species tend to avoid extinction longer than specialized species. Species, however, are not individuated solely in terms of their genotypic or phenotypic properties. If a group of organisms were to evolve that was identical in every respect except ancestry to the now extinct *pterodactyl*, it would be recognized as a distinct species. Other examples are not hard to find.¹⁹ Linguistic theory, as we have seen, can type-distinguish two physically identical forms on the basis of their contexts of utterance. Claims that geology, biology, or linguistics are *unscientific*, or the processes they describe not *causal*, because they employ individuating principles which take account of features of the history or context of the states and objects they describe, are simply unwarranted. The taxonomic schemes of science are responsive to other considerations besides micro-structure. Fodor would surely agree that psychology should not be held to stricter standards than these other sciences.

Fodor's "metaphysical point," then, is not supported by scientific practice. Moreover, the intelligibility of mental causation does not require that mental states supervene on local micro-structure, but only that mental states be physically realizable as brain states.²⁰ Supervenience, after all, is a relation between *types*, whereas causation is presumably a relation between state/event *tokens*. For mental causation to be possible there must be some mechanism connecting token mental states with token behavioral states.²¹ Fodor has offered no cogent reason to think that mental states individuated with respect to non-locally supervenient (that is, "wide") contents cannot be physically realized as brain states.

II. EMPIRICAL EVIDENCE

The question whether psychology is or should be individualistic cannot, I believe, be settled by appeal to general metaphysical or

¹⁹Burge gives other examples in "Individualism and Psychology," p. 16.

²⁰Whether mental causation requires supervenience on physical states of the individual *and* the environment (including the linguistic community) is another matter, beyond the scope of this paper.

²¹The worry here is that individuating mental states widely requires postulating mysterious causal mechanisms. I have argued that it does not. This question is quite distinct from the worry that the causal efficacy of mental states, *qua mental*, is inconsistent with the causal closure of the

methodological principles. Perhaps, though, empirical considerations might be brought to bear. In "Individualism and Psychology," Tyler Burge offers two arguments, purportedly based on psychological practice, for the claim that a substantial portion of psychology, as currently practiced, is non-individualistic. Neither argument, I shall argue, supports Burge's claim.

The first line of argument appeals to the Putnam/Burge thought experiments. Let me briefly sketch Burge's own example. He asks us to consider an individual, Burt, who believes that he has arthritis in his thigh, not knowing that arthritis is a disease of the joints. His molecular duplicate, Twin Burt, lives in a community where the term "arthritis" is used to refer to a disease which afflicts not only the joints but the long bones of the legs as well. Whereas both subjects would express their beliefs using the form of words "I have arthritis in my thigh," the contents expressed by their words would seem to be different, reflecting the difference in their linguistic communities. Only Burt has a belief about arthritis. Or so Burge argues.

The argument that individualism is false of psychology, as it is currently practiced, proceeds as follows: psychology attributes intentional mental states to agents using interpreted that-clauses (so-called oblique attitude ascriptions, for example, the belief *that water is wet*). The content attributed by such that-clauses does not supervene on non-intentionally specifiable states of the subject, as is shown by the thought experiments. Since there is no reason to assume that the attribution of content by psychological theory is instrumentalist or merely heuristic, Burge concludes that this part of psychology, at least, is non-individualistic.

Burge's argument is inconclusive. Even granting (1) that the thought experiments do show that oblique (*de dicto*) attitude ascriptions are sensitive to environmental and social factors, and (2) that psychology makes use of interpreted that-clauses in ascribing propositional attitudes, recent arguments by Brian Loar²² suggest that these two premises alone do not imply that psychology indi-

physical world, and the non-reducibility of mental properties to physical properties. This issue is orthogonal to the present dispute.

²²Brian Loar, "Social Content and Psychological Content," in *Contents of Thought*, ed. R. Grimm and D. Merrill (Tucson, Ariz.: The University of Arizona Press, 1988), pp. 99–110.

viduates mental states partly by reference to environmental or social factors. According to Loar, the content of mental states, as individuated by common-sense psychological explanation (what he calls *psychological content*), is not, in general, to be identified with what is picked out by oblique attitude ascriptions (so-called *social content*). Sameness of *de dicto* or oblique attitude ascription, he claims, is neither necessary nor sufficient for sameness of psychological content.

The claim that sameness of oblique attitude ascription is not *necessary* for sameness of psychological content challenges the conclusion Burge draws from the thought experiments. The beliefs of the two protagonists in the examples are ascribed using different oblique attitude ascriptions (whereas I believe *that water is wet*, my twin believes *that twater is wet* (or some such); Burt believes (incorrectly) *that he has arthritis in his thigh*, whereas Twin Burt believes (truly) *that he has tharthrititis in his thigh*). It is alleged to follow from this fact that the two individuals have different beliefs.

Loar asks us to consider a diary, written by either an earthling or a twin-earthling, although we do not know which. One entry says: "No swimming today; the water is too rough." As common-sense psychologists we would explain the diarist's behavior as follows: because she believes that the water is rough, and believes also that if the water is rough one should not swim, she believes that one should not swim. If our explanation is challenged on the grounds that we don't know whether she has any beliefs about *water* at all, then, according to Loar, we would likely produce a paraphrase of the original explanation that makes no commitment about the kind of liquid the diarist is talking about (for example, "She believes that the local sample of the liquid that fills the oceans and lakes is rough . . ."); that is, some *single* explanation that would allow us to subsume the beliefs of both earthlings and twin-earthlings, rather than *two* explanations, one adverting to water and the other to twater. On Loar's view, our original psychological explanation attributes to the diarist only a certain way of conceiving of things; in particular, it attributes mental states that are interrelated (and related to stimuli and behavior) in specified ways, but involves no commitment about the environmental or social context in which the beliefs are embedded.

According to Loar, the reported psychological explanation loses nothing from the fact that we do not know whether the diarist is

talking about water or some other liquid. We understand the psychological explanation, Loar claims, because we know how the diarist personally conceives things. Whether she is an earthling or a twin-earthling, the psychological content of her mental state is the same.

There is something right about this. The explanation, *qua* psychological explanation, seems both correct and *complete*. We can make further predictions of the diarist's behavior without knowing whether the liquid that she is considering swimming in is H₂O or XYZ. The latter issue just seems psychologically irrelevant. If we explain the diarist's behavior as Burge would have us do—by constructing two explanations—then a psychological generalization would appear to be lost.

According to Loar, the feature of beliefs that determines their type-individuation by common-sense psychology is their pattern of actual and potential interaction with other beliefs, with stimuli, and with behavior, that is, their conceptual (functional) role. Only as individuated by their conceptual roles, Loar argues, are mental states subsumed under the generalizations that constitute psychological theory. One cannot simply assume, as Burge seems to, that the appropriate individuating conditions are captured by oblique attitude ascriptions.

A second example, designed to show that sameness of oblique attitude ascription is not sufficient for sameness of psychological content, makes the above claim plausible. Loar describes an English speaker, Paul, who believes that he has arthritis in his ankles. Paul visits Paris and hears of a disease called "arthritis." Not realizing that "arthritis" and "arthrite" are intertranslatable, and being a bit of a hypochondriac, Paul comes to believe that he has two problems with his ankles, one he calls "arthritis," the other "arthrite." Loar concludes that Paul has *two* beliefs concerning his ankles, each with distinct psychological contents, despite the fact that only one oblique attitude ascription is true of him (namely, he believes *that he has arthritis in his ankles*).

The ground for attributing to Paul two type-distinct beliefs about his ankles is that his "French" belief and his "English" belief have different interactive potential—that is, they have different causal powers. Paul's "French" belief that he has arthritis in his ankles does not interact in the appropriate way with his "English" belief that if he has arthritis in his ankles he should take aspirin,

although his “English” belief that he has arthritis in his ankles does. The explanatory schemas of common-sense psychology subsume Paul’s mental states only as they are individuated by their conceptual roles.

It might be objected that in virtue of Paul’s two beliefs about his ankles *two* relevant oblique attitude ascriptions are true of him—perhaps (1) he believes *that he has an ailment called “arthritis” in his ankles*, and (2) he believes *that he has an ailment called “arthritis” in his ankles*. But Paul’s beliefs about his ankles do not seem to be meta-linguistic beliefs; moreover, it is difficult to find alternative candidates for the relevant attitude ascriptions. (*That he has arthritis in his ankles* is not well formed in any language and hence not a candidate.) The objection seems to rest on the assumption that distinct beliefs *require* distinct attitude ascriptions, but of course, this claim is precisely what is in dispute.

If Loar’s analysis of the examples is correct, then he seems to have identified a kind of “narrow” content which plays a prominent role in common-sense psychological explanation. In response to the objection that there is no way to *specify* narrow contents, since that-clauses are shot through with environmental and social presuppositions, Loar argues that we are able to individuate mental states for the purpose of psychological explanation without a means of specifying psychological content directly. We get at them in context through various devices; for example, we report a person’s words along with other utterances or facts that help us to interpret her words, as in the description of the examples. If twin cases (that is, individuals identical, or more plausibly, similar in physical makeup or functional organization) were to become common then psychology would probably develop precise, more convenient ways of picking out mental states (or their conceptual roles), but for now, outside the philosopher’s laboratory, that-clauses serve well enough.

It is not my concern to argue that Loar is correct in distinguishing two types of content. The point I want to make is a more general one. Properties of belief *ascriptions* (that-clauses) should not be attributed without further argument to the mental states that they serve to pick out in psychological theory, especially since it is a widely acknowledged fact that belief ascriptions serve purposes in everyday discourse besides predicting and explaining the behavior and cognitive capacities of the subjects to whom they are

ascribed. Loar's examples suggest the possibility that psychological theory may individuate mental states individualistically even though the language it uses to refer to these states is not itself individualistic. The thought experiments themselves, then, and the fact that psychology uses that-clauses in ascribing intentional content, do not establish that psychology is non-individualistic in its taxonomic practices.

* * *

Burge's second line of argument appeals to David Marr's²³ theory of early vision in support of the claim that individuating practices in psychology involve presuppositions about the specific nature of the subject's environment. Perceptual states are alleged to be individuated in Marr's theory by reference to the distal stimulus which gives rise to the image that forms the input to the visual system.

Burge argues as follows:²⁴

- (1) Marr's theory is intentional.
- (2) The intentional primitives of the theory and the information they carry are individuated by reference to contingently existing physical items or conditions by which they are normally caused and to which they normally apply.
- (3) So if these physical conditions and, possibly, attendant physical laws were regularly different, the information conveyed to the subject and the intentional content of his or her visual representations would be different.
- (4) We can conceive of relevantly different physical conditions and perhaps relevantly different (say, optical) laws regularly causing the same non-intentional, individualistically individuated physical regularities in the subject's eyes and nervous system.
- (5) In such a case (by 3) the subject's visual representations would carry different information and have different rep-

²³David Marr, *Vision* (New York, N.Y.: Freeman and Co., 1982).

²⁴"Individualism and Psychology," p. 34.

representational content though the person's whole non-intentional physical history (at least up to a certain time) might remain the same.

- (6) The theory individuates some perceptual states in terms of their informational or intentional content.
- (7) Therefore, individualism is false for the theory of vision.

Step (1), Burge claims, is "evident" from the fact that the "top levels" of the theory are formulated in intentional terms. Step (2), he claims, is supported by a number of examples from the theory, one of which describes how general physical assumptions constrain the choice of the representational primitives out of which the so-called primal sketch²⁵ is constructed. Step (3) is said to be supported by the above considerations regarding the individuation of intentional content, and by examples illustrating the explanatory method of the theory.²⁶ Step (4) constructs a Twin-Earth-type thought experiment. Step (5) follows from steps (3) and (4). Step (6) is claimed to be "unproblematic."

Burge's argument may be challenged at a number of points,²⁷ but I shall focus on the claim (step 1) that Marr's theory is intentional. I shall argue that Burge's confidence in this claim is unfounded. In order to develop the point, I begin with a brief sketch of Marr's theory.

The theory of vision, as conceived by Marr, is deployed at three different levels of description. There is the specification of *the*

²⁵Marr defines the primal sketch as "a representation of the two-dimensional image that makes explicit the amount and disposition of the intensity changes there" (*Vision*, p. 366).

²⁶One example cites the role of physical constraints in visual processing, in particular, the *constraints on matching* that facilitate stereopsis. According to Burge:

The methods of individuation and explanation are governed by the assumption that the subject has adapted to his or her environment sufficiently to obtain veridical information from it under normal conditions. If the properties and relations that *normally* caused visual impressions were regularly different from what they are, the individual would obtain different information and have visual experiences with different intentional content ("Individualism and Psychology," p. 33).

²⁷Gabriel Segal, for example, attacks premise (2) of Burge's argument in "Seeing What is Not There," *The Philosophical Review* 98 (1989), pp. 189–214 (see n. 35 below).

function computed by the various visual processes hypothesized by the theory, *the algorithmic implementation* of the hypothesized functions, and *the hardware implementation* of the hypothesized algorithms. Marr stresses the importance of the first of these levels—what he calls the “theory of the computation”—arguing that the nature of the computational processes underlying vision depends more upon the problems to be solved than upon the hardware in which the solutions are realized.

The goal of the visual system is to derive a representation of three-dimensional shape from information contained in two-dimensional images. Marr’s theory divides this task into four distinct representational stages; corresponding to each is a type of representation, tokens of which serve as the inputs and/or outputs to the computational processes that effect the derivation of shape information from the image.

The computational processes to which tokens of these representations serve as inputs and/or outputs, up to the construction of what Marr calls the *2.5-D sketch*,²⁸ are data-driven (“bottom-up”) processes which rely only on information contained in the image. This is accomplished despite the fact that the surface description of the distal scene is underdetermined by such information, because the visual system incorporates certain very general *assumptions* about the environment that reflect physical constraints on the pairing of retinal images with distal shapes. These innate assumptions provide the additional information necessary to recover shape from the information contained in the image. An example of such an assumption is Shimon Ullman’s *rigidity assumption*, which says that “any set of elements undergoing a two-dimensional transformation has a unique interpretation as a rigid body moving in space and hence should be interpreted as such a body in motion.”²⁹ Ullman has proved that three distinct orthographic views of four non-coplanar points in a rigid body are sufficient to determine its three-dimensional structure (the structure from motion theorem). In a world like ours where most things are rigid, a pro-

²⁸The 2.5-D sketch is a viewer-centered representation of the depth and orientation of visible surfaces, including contours of discontinuity in depth and orientation.

²⁹Shimon Ullman, *The Interpretation of Visual Motion* (Cambridge, Mass.: The MIT Press, 1979), p. 146.

cess that incorporates the rigidity assumption would generally be successful in recovering the three-dimensional structure of distal objects from three such views. Other processes incorporate similar general assumptions to recover information about shape from properties of the image.³⁰

I return now to Burge's argument, in particular, to the claim at step (1) that Marr's theory is intentional. Burge takes this claim to be "evident" from the supposed fact that the "top levels" of the theory are formulated in intentional terms. The claim, however, is far from evident. While it is true that most of the inputs and outputs of the various computational processes postulated by the theory are given an intentional *interpretation* in Marr's exposition of these processes, it is by no means obvious that these representational tokens are presumed *by the theory* to have a distal, and hence intentional, interpretation. The tokened primitives of the primal sketch, for example, are not, as Burge claims (p. 30), identifiable aspects of the distal scene. Rather, as Marr emphasizes repeatedly, they are abstract properties of the retinal image—blobs, terminations, virtual lines, etc., and aggregations of such items:

[T]he primitives [of the primal sketch] can be defined from the image in a variety of ways, from the very concrete (a black mark) to the very abstract (a cloud of dots).³¹

[These primitives are] descriptions of an image at different scales.³²

The primal sketch is a very rich description of an image.³³

I should note in passing that the processes responsible for stereopsis and directional selectivity take as inputs so-called zero-crossings³⁴ from individual channels. Such inputs, Marr points out, are *not* "physically meaningful"—that is, they have no specifi-

³⁰It should be noted that Marr's theory involves no commitment about how these assumptions are incorporated in the system. In particular, there is no claim that they are explicitly represented.

³¹*Vision*, p. 52.

³²*Ibid.*, p. 53.

³³*Ibid.*, p. 71.

³⁴A zero-crossing is a point where the value of a function changes its sign. Zero-crossings correspond to sudden intensity changes in the image.

able distal interpretation, and hence are clearly not intentional primitives.

But even if it were true that all of the inputs and outputs of these processes did receive an intentional specification in Marr's articulation of the theory, it would not follow that the theory was intentional. The crucial issue that Burge fails to address is whether in fact these intentional specifications are part of the theory. Burge assumes that everything that Marr says by way of a description of the hypothesized visual processes counts as part of the theory, but the assumption is dubious. Just as in the exposition of theories in other sciences (physics, chemistry, etc.), where it is crucial to distinguish the theory from its various models and interpretations, it is crucial here to distinguish the computational *theory* from the attendant expository apparatus, specifically the intentional interpretation of the computational primitives that renders the theory intelligible.

The task of delimiting Marr's theory of early vision is made especially difficult by the fact, noted above, that his theory is articulated at three different levels of description. Burge assumes that the first, or "top-most," of these levels—what Marr calls "the theory of the computation"—is intentional, because Marr typically provides a rough intentional specification of the function computed. Yet Marr's provision of such a specification hardly makes the theory intentional in the sense that Burge's argument requires. To see this, consider the sort of case envisioned by Burge in step (4) of his argument. For example, consider a subject for whom Marr's theory provides a correct account of her visual processes. Suppose this individual were transported into an environment in which her intentional perceptual states would, on Burge's account, have different content. On Burge's account, Marr's theory of early vision true of such a subject would have changed, since the content of the intentional primitives that supposedly figure in the theory of the computation would have changed. This seems implausible. Marr's theory of early vision is clearly not intended as a theory true of its subjects only in this environment.

That Marr's theory should remain true of its subjects in other environments does not, of course, entail that the theory is non-intentional. It might be argued that Marr's theory does contain intentional primitives, but that the content ascribed to the postulated representational tokens in the theory is "narrow," that is, in-

variant across the environmental changes envisioned in steps (3) and (4) of Burge's argument. In other words, one might simply deny premise (2) of Burge's argument, the claim that the intentional primitives of the theory are individuated by reference to their normal distal causes in this environment.³⁵ The point is rather that the stated goal of Marr's theory is to characterize the computational *mechanisms* that underlie our visual capacities. These mechanisms will not change in the imagined counterfactual environments. The functions computed by the various visual processes—the so-called theory of the computation—will not change; the algorithms that implement these functions will not change; nor will the hardware that implements these algorithms. The only thing that might change would be the intentional *specification* of the functions computed.

If, as I have argued, the same functions are computed across different environments, then why should the specification of these functions change from one environment to the next? Why does Marr not simply specify the function in a way that is neutral with respect to the particular environment in which the organism is situated—for example, as a function on the natural numbers? There are at least three reasons why Marr might have chosen to provide intentional specifications of the functions computed by the visual system. In part, I think, he did so for the same reason that Bohr presented his theory of the atom by way of relatively familiar notions—because the postulates of the theory can be understood more readily when embedded in a model than when given a more complicated purely formal exposition. An intentional specification of the functions computed by the visual system (more specifically, of the representational tokens over which the computational pro-

³⁵Gabriel Segal defends just this position in "Seeing What is Not There." But the postulation of narrow content does not appear to be justified by Marr's own methodology. Such "non-specific" contents would seemingly have to be wildly disjunctive, containing a disjunct for every possible "normal cause" of an input. As Segal points out, "what is represented at a given stage is tightly constrained by what is exploited at the next stage" (p. 195). It is hard to imagine that such contents could play an explanatory role in subsequent computational processing, in particular, a role that could not be played by an abstract property of the image itself (that is, an uninterpreted token). So the attribution of narrow content in the theory does not appear to be well motivated.

cesses of the theory are defined) makes these processes intelligible *in the context in which the subject is situated*.

Second, and perhaps more importantly, the use of intentional descriptions enables us to explain how the visual system functions in the environment in which the subject is located. It is a plausible adequacy condition on a theory of vision that it enable us to understand how vision contributes to a subject's success (or failure) in its interactions with its environment. Marr's theory meets this adequacy condition in part by specifying the functions computed by various visual processes by reference to general features of the subject's environment. That is to say, the theory specifies the computational processes that form the visual system in a way that allows us to see how they are able to contribute to their possessor's successful (or unsuccessful) interactions with her environment. By employing an intentional specification of the representational tokens over which the hypothesized computational processes of the visual system are defined, Marr is able to show that in an environment like our own these processes succeed in recovering information about that environment. But precisely because the hypothesized visual mechanisms are invariant across environmental changes, we can also see why *this* visual system would *not* be adaptive had the environment been different.

Finally, given that the questions that define a psychological theory's domain are often couched in intentional terms (for example, a theory of vision must explain how the visual system derives a representation of three-dimensional shape of the distal scene from information contained in two-dimensional images), one rationale for articulating the theory in intentional terms is to demonstrate that these questions receive an answer in the theory (when combined with certain assumptions about the organism's environment).

But the crucial point remains: one must be careful not to mistake an adventitious feature of a particular model of a theory for an essential feature of the theory itself. This is precisely what Burge seems to have done. If the foregoing account is correct, then Marr's theory of vision is not intentional, although it does have intentional models. The primitives in the theory are not intentional, although some of the primitives do have intentional specifications. I am inclined to think that the foregoing is correct;

however, for present purposes it is enough to show that it is not at all "evident," as Burge claims, that Marr's theory is intentional, simply because the "top levels" of the theory are formulated in intentional terms. Burge's entire argument collapses with this first step, since if the theory is not demonstrably intentional, any argument to the effect that the theory is non-individualistic is a non-starter. The question whether Marr's theory is intentional is more complex than Burge seemingly realizes. Until we have a much more sophisticated philosophy of psychology, which is able to delimit a psychological theory from the expository apparatus that renders it intelligible, bringing psychological practice to bear on questions of individualism is not likely to be fruitful.

III. CONCLUDING REMARKS

We have seen that Fodor's metaphysical argument that psychology must be individualistic doesn't go through; nor do Burge's reflections on various aspects of psychological theorizing establish that individualism is false. Philosophers of science not caught up in the dispute regarding individualism might reasonably have supposed that the question of whether psychology is or should be individualistic turns crucially on the *goals* of psychological theorizing, something that both Fodor and Burge largely neglect. If the goal of a theory is to characterize the mechanisms underlying some cognitive capacity, then we might well expect the theory to employ an individualistic taxonomy. This would be especially true if, as in the case of Marr's theory of vision, we are concerned to understand the mechanism's fitness for this and other environments. If, on the other hand, the theory takes cognitive capacities as given, and tries to give an account of a subject's behavior in a particular environment, as social psychology does, then the theory is unlikely to be individualistic. For in such cases the concern is with the behavior of an organism embedded in a context, and non-individualistic characterization of the subject's mental states would impose a matrix of socially determinate relations on individual psychology.

So, to the question, "Must psychology be individualistic?" the correct answer is presumably this: No, psychology does not have to

MUST PSYCHOLOGY BE INDIVIDUALISTIC?

be individualistic. Sometimes it is, sometimes it isn't. It all depends on the focus or goals of the theorizing. Computational theories such as Marr's presumably are individualistic inasmuch as the stated goal of such theories is to characterize the computational mechanisms that underlie various cognitive capacities.³⁶

Rutgers University

³⁶Earlier versions of this paper were presented at Wayne State University, Rice University, The University of Notre Dame, The University of Calgary, and Rutgers University. My thanks to those who participated in the discussions. I am also indebted to Robert Allen, Denise Gamble, William Harper, Ausonio Marras, and especially Robert Matthews for helpful comments and criticism.