The ascendency of functionalism marks a period of optimism in the philosophy of mind. In recent years, functionalism has provided the dominant expression of the hope that beliefs, desires and intentions, as attributed by means of 'that'-clauses, can be accommodated by a scientific theory of mental processes. I shall regard as functionalist those views which define psychological states in terms of causal relations among sensory inputs, internal states and behavioral outputs, all ultimately describable in terms applied to automata; and my claim will be that attitudes, such as beliefs¹ identified by 'that'-clauses, can not be construed as functional states, on pain of contradiction.

Classical functionalism² fails because it is caught in an unrecognized dilemma, a dilemma concerning the individuation of psychological states that explain behavior. Beliefs are individuated by most functionalists in terms of that 'that'-clauses; functional states are individuated 'narrowly' (i.e., specifiable without presupposing the existence of anything other than the individual whose states they are). If beliefs are to be functional states, individuation in terms of 'that'-clauses (on some construal) must coincide with genuinely 'narrow' individuation. Items that can not be identified in terms of 'that'-clauses do not qualify as beliefs; items that can not be identified narrowly do not qualify as functional states. But individuation in terms of the narrowest of 'that'-clauses (i.e., in terms of narrow semantic type) leads to the following dilemma: Either the functionalist is committed to an inconsistent triad, or no states identified by 'that'-clauses are sufficiently narrow to allow beliefs to be functional states.

ON BEING NARROW

Beliefs are usually identified by functionalists and others in terms of the 'that'-clauses of their ascriptions: belief that $p$ differs from belief that $q$ in

---

virtue of differences between \textit{that} \( p \) and \textit{that} \( q \). However, controversy thickens around specification of the relevant differences between \textit{that} \( p \) and \textit{that} \( q \) that constitute the difference between believing \textit{that} \( p \) and believing \textit{that} \( q \).

The functionalist claim is that the beliefs that play a role in the etiology of (brute) behavior must be understood as functional states. Functionalists recognize that many beliefs — indexical beliefs, \textit{de re} beliefs — are not plausibly construed as functional states, and hence fall outside the purview of the functionalist claim. Since the functionalist concern is with beliefs and other psychological states that explain behavior, the issue becomes one of construing ‘\textit{that}’-clauses (‘\textit{contents}’) in such a way that behavior-explaining beliefs can be understood as functional states.

If beliefs are to be understood as classical functional states, they must be construed as \textit{narrow states}. As the idea of narrowness was introduced by Putnam, psychological states in the narrow sense were to be those permitted by methodological solipsism, that is, by the “assumption that no psychological state, properly so-called, presupposes the existence of any individual other than the subject to whom that state is ascribed”\(^3\). Fodor brought the idea of \textit{narrow content} to prominence as the constraint that mental states that explain behavior should be individuated by content without regard to truth or reference to individuals, or as the constraint (which may not be the same thing) that for purposes of explaining behavior, attitudes should be attributed “opaquely” rather than “transparently”\(^4\). Yet another characterization of narrowness that has currency is that narrow states are those shared by molecule-for molecule duplicates.

It may be thought that the descriptions ‘state common to molecular replicas’, ‘state specifiable without regard to truth or reference’, ‘state specifiable without presupposing the existence of any individual other than the subject to whom that state is ascribed’ and ‘state “opaquely” attributed’ are co-extensive. Such an assumption is abetted by the widespread view that there is an interesting sense of ‘behavior’, which requires explanation solely in terms of what is “in the head” of the subject. Although all the criteria for narrowness aim to isolate psychological states that are “in the head”, it does not follow that they in fact isolate the same things; indeed, I am confident that they do not.\(^5\) However, no matter how narrow states are characterized, every criterion rules out beliefs individuated in part by objects in the believer’s environment as “wide”\(^6\).
Let us say: A state is narrow if and only if whether or not $x$ is in that state is determined solely by properties of $x$, without presupposing that anything other than $x$ exists. The belief that $p$ has narrow content if and only if whether or not $x$ believes that $p$ is determined solely by properties of $x$, without presupposing that anything other than $x$ exists. Classical functional states are narrow; and if functionalism is correct, certain beliefs also have narrow content — i.e., beliefs that explain (brute) behavior may be individuated without presupposing that anything other than the believer exists.

I shall use the locution ‘narrow semantic type’ to indicate beliefs individuated by ‘that’-clauses in a restricted sense: Ordinary individuation of beliefs is by means of the obliquely-occurring expressions in ‘that’-clauses of ascriptions, where obliquely-occurring expressions are those for which substitution of co-extensive terms is impermissible. Criterion (I), which individuates beliefs by obliquely-occurring expressions in content clauses, suggests another formulation, criterion (II): Belief that $p$ is a different belief from belief that $q$ if and only if there are semantic differences between $p$ and $q$ other than differences in truth-value or reference to specific individuals. Although (I) and (II) are nonequivalent, for purposes here, I shall assume that beliefs that are different on (I) are also different on (II), and that beliefs individuated by either (I) or (II) are individuated by narrow semantic type.

So, belief tokens that differ only in truth-value (e.g., two tokens of the belief that the drinking age in Vermont is eighteen, one in 1984 and the other in 1986) or in objects denoted (e.g., two tokens of the belief that the U.S. President is a Democrat, one in 1942 and the other in 1946) may be of the same narrow semantic types. Beliefs not typically identified by narrow semantic type include de re beliefs, indexical beliefs, beliefs in singular propositions, and perhaps beliefs expressed by proper names. In general, individuation of beliefs partly by objects external to the believer will not coincide with individuation of beliefs by obliquely-occurring expressions in ‘that’-clauses of ascription (i.e., by narrow semantic type). Examples of beliefs that differ in narrow semantic type are the beliefs that fishing is fun and that hiking is healthful.

The notion of narrow semantic type is intended only to capture a certain aspect of the intuitive identification of belief via ‘that’-clauses. Since I make no theoretical use of the notion, it can be adjusted for various semantical views. In particular, my use of ‘narrow semantic type’ leaves it open whether or not differences in narrow semantic type amount to any more than differences in truth-condition (as opposed to truth-value). The argument leading
to the second horn of the dilemma, for example, rests on the claim that while beliefs may be individuated (as wide states) by truth-conditions or (as genuinely narrow states) by functional role, there is no place for narrow semantic type as a tertium quid; all differences in narrow semantic type can be handled as differences in truth-conditions. Whatever the merit of this position, as we shall see, it can not save the functionalist construal of beliefs as functional states.

Here, then, is the key question: does individuation by narrow semantic type coincide with individuation by functional state? An affirmative answer leads to the inconsistent triad (the first horn); a negative answer leads to a denial that beliefs, individuated by obliquely-occurring expressions in ‘that’-clauses (i.e., by narrow semantic type), are functional states (the second horn).

The inconsistent triad is this:

(A) Beliefs individuated by narrow semantic type are psychological states.
(B) Psychological states are functional states.
(C) Two tokens of a single functional type may differ in narrow semantic type.

The reasoning leading to commitment to (A)—(C) (the first horn of the dilemma) is found in Fodor’s ‘Methodological solipsism considered as a research strategy in cognitive science’. The only way that I know for a classical functionalist to avoid the contradiction leads straight to the second horn of the dilemma. The argument out of the inconsistent triad consists of maintaining that individuation by narrow semantic type requires recourse to truth-conditions and hence that individuation by narrow semantic type does not individuate beliefs narrowly enough to coincide with individuation by functional state. I shall give an example of this approach, suggested in several articles by William G. Lycan, and then point out that the consequence is a denial that beliefs are functional states.

**FIRST HORN: AN INCONSISTENT TRIAD**

(A) The term ‘narrow semantic type’, is intended to accommodate an intuitive sense in which beliefs are said to have ‘content’. As Jerry Fodor illustrates the notion of content, to think that Marvin is melancholy is to be in
a particular relation to a "representation the content of which is that Marvin is melancholy". And the belief that Marvin is melancholy is distinguished from the belief that it sometimes shows in Cincinnati by the contents of the associated representations. Fodor's initial assumption is that "the content of representation is a (type) individuating feature of mental states". [MS, 63, 64; emphasis his]

Fodor argues that (at least some) psychological states that explain behavior are to be understood in terms of "opaque" attribution of attitudes:

Thesis: when we articulate the generalizations in virtue of which behavior is contingent upon mental states, it is typically an opaque construal of the mental state attributions that does the work; for example, it's a construal under which believing that a is F is logically independent from believing that b is F, even in the case where a = b. It will be convenient to speak not only of opaque construals of propositional attitude ascriptions, but also of opaque taxonomies of mental state types.... So, the claim is that mental states are typically opaquely taxonomized for purpose of psychological theory. [MS, 66; emphases his.]

This suggests that it is beliefs individuated by narrow semantic type that are the psychological states that explain behavior. Thus, (A) is one of Fodor's central claims.9

(B) is the backbone of functionalism. To say that psychological states are functional states is to say that, for purposes of psychological classification, mental tokens are typed nonsemantically, by their (typical) causes and effects. Long a champion of functionalism, Fodor has elaborated a "formality condition", according to which mental processes apply to representations in virtue of their formal or nonsemantic (e.g., syntactic or computational) properties, in support of functionalism. [MS, 64]

Since content is a "semantic notion par excellence", [MS, 64] there may seem to be a conflict between (A), according to which certain mental tokens are classified by narrow content (understood as coinciding with narrow semantic type), and (B), according to which those mental tokens are classified by their nonsemantic properties. But Fodor takes it as a desideratum of a theory of the mind that it reconcile semantic and nonsemantic properties of mental tokens:

By thus exploiting the notions of content and computations together, a cognitive theory seeks to connect the intensional properties of mental states with their causal properties vis-à-vis behavior. Which is, of course, exactly what a theory of the mind ought to do. [MS, 68]

Such congruence is supported by the claim (or conjecture)
... that mental states are distinct in content only if they are relations to formally distinct mental representations; in effect, that aspects of content can be reconstructed as aspects of form, at least insofar as appeals to content figure in accounts of the mental causation of behavior. [MS, 68]

The argument for this claim is that it provides an explanation of “how beliefs of different content can have different behavioral effects”.

The form of explanation goes: it’s because different content implies formally distinct internal representations (via the formality condition) and formally distinct internal representations can be functionally different; can differ in their causal role. Whereas, to put it mildly, it’s hard to see how internal representations could differ in causal role unless they differed in form. [MS, 68]

Thus, Fodor is unlikely to take exception to (B). In any case, although (B) is stated as a specifically functionalist thesis, other ways of typing mental tokens nonsemantically (in conformity with the formality condition) can play the same role in generating the contradiction below.

(C) also seems acknowledged by Fodor, who explicitly considers computer programs that simulate, on the one hand, the Six-Day War and on the other hand, a chess game: “It’s a possible (though, of course, unlikely) accident that these programs should be indistinguishable when compiled; viz., that the [machine language] counterparts of these programs should be identical, so that the internal career of a machine running one program would be identical, step by step, to that of a machine running the other”. In addition: “machines typically don’t know (or care) what the programs that they run are about; all they know (or care about) is how to run their programs. This may sound cryptical or even mystical. It’s not. It’s merely banal”.

I believe that Fodor here underestimates the significance of his own observation. As Georges Rey has remarked even more pointedly:

[O]n Wednesday [a computer] deals with the intricacies of the SALT negotiations, on Thursday it plays chess with Bobby Fischer. Now it is perfectly possible in principle for the machine to pass through type identical computational and physical states on both days. All that would normally be needed is that on both occasions the input decks be themselves type identical, and that would occur should the two problem domains be construed, as it were, isomorphically. It’s just that on Wednesday the punches in the cards are interpreted (say, by Carter) to refer to Brezhnev, Vienna and 100-megaton bombs; and on Thursday the very same punches are interpreted (say, by Spassky) to refer to moves and pieces in chess.

So although a belief that cruise missiles can repel an invasion from the East, say, is distinct in narrow content from a belief that pawns can ward off an attack on the Queen’s side, insofar as a computer may be said to have beliefs
at all, both beliefs may be equally "subserved" by a single type of computational and physical state. Then the computer, in exactly the same sequence of functional states, could equally well be interpreted as reasoning about negotiating an arms agreement or about winning a chess game. Thus, tokens of a given functional type may be of different narrow semantic types. Indeed, the fact that a token of a particular functional type is of one narrow semantic type does not preclude its also being of another (nonequivalent) narrow semantic type. So, I take it both that (C) is true and that Fodor is committed to it.

The three theses together lead to a contradiction. In case the contradiction is not apparent, let me provide a simple proof. 'E' will stand for the existential quantifier, and universal quantifiers will be omitted.

Let $x, y$ range over narrow semantic states.\(^{12}\)

\[ P_x : x \text{ is psychological state No. 34.} \]
\[ F_x : x \text{ is functional state No. 7.} \]

$a, b$ are narrow semantic state constants.

Suppose that a person is in a narrow semantic state $S_I$ if and only if the person has a belief individuated by narrow semantic type $T_I$. Then, according to (A), distinct narrow semantic states are distinct psychological states, in which case if one narrow semantic state is psychological state No. 34, then a different narrow semantic state is not psychological state No. 34.

\[
(1) \quad (x \neq y) \supset (P_x \supset \sim P_y)
\]

According to (B), distinct psychological states are distinct functional states, in which case if a semantic state's being psychological state No. 34 precludes a second semantic state's also being psychological state No. 34, then the first semantic state's being functional state No. 7 precludes the second semantic state's being functional state No. 7.

\[
(2) \quad (P_x \supset \sim P_y) \supset (F_x \supset \sim F_y)
\]

According to (C), distinct narrow semantic states may be a single functional state. To simplify the argument, assume that there is an instantiation of this possibility.

\[
(3) \quad (Ex, y)(x \neq y \& F_x \& F_y)
\]
Instantiate 3:

(4) \[ a \neq b \land Fa \land Fb \]

Instantiate (1) and (2) apply the rule for hypothetical syllogism:

(5) \[ (a \neq b) \supset (Fa \supset \sim Fb) \]

From (4) and (5), via modus ponens derive:

(6) \[ Fa \supset \sim Fb \]

After applying the modus ponens rule to (6) and (4), conjoin the consequent of (6) with the last conjunct of (4):

(7) \[ Fb \land \sim Fb. \]

Notice that it is no help to make explicit relativity of narrow semantic type to an overall scheme of interpretation. If narrow semantic types were relativized to interpretations, and new variables, \( I_i \), etc., were introduced to range over interpretations, evidently the argument would still go through. Moreover, although states of a machine may be given any consistent interpretation that serves the purposes of the user or programmer, there is no such license in attributing attitudes to people. Moral considerations aside, it would be an outright error to attribute to Carter beliefs concerning moves in chess when he was negotiating with Brezhnev, regardless of his internal states.

The significance of the first horn of the dilemma is this: As long as functional state tokens can have more than one semantic interpretation — and I take it to be a central feature of the computer analogy that they can — mental states like beliefs can not be understood as functional states. So unless Fodor and other functionalists are willing to abandon the view that beliefs are functional states (the second horn), I do not see how they can avoid the contradiction, which stems from the very machine analogy that has given functionalism its impetus.

How is it that an easily derivable contradiction has been overlooked? I would conjecture that the answer lies in the apparent fruitfulness of the machine analogy itself. Like any analogy, the machine analogy has two moments, as it were. It must draw both on features of computers and on features of mental states. None of the features relevant to the contradiction has gone unremarked by functionalists; nor, as we have seen, have the consequences of holding (A) and (B) together been overlooked. What no one has
thought to do is to conjoin the "mind side" of the analogy (as represented by (A) and (B)) with the "machine side" of the analogy (as represented by (C)). It is only here that the contradiction emerges.

On the one hand, if psychological states are functional states, it seems almost evident that there can be no semantic difference between psychological states without a functional difference between them. On the other hand, implicit in the computer model is the idea that any functional state may be given more than one interpretation, and since interpretation is "a semantic notion par excellence" [MS, 64], there can be differences in narrow semantic type without functional differences. So the possibility even of molecular duplicates with different beliefs individuated by narrow semantic type is implicit in the computer analogy itself. Thus, it seems that two aspects of the computer analogy are in conflict, as long as the analogy is taken to apply to mental states, such as beliefs, desires and intentions, which are identified in part by narrow semantic type.13

SECOND HORN: BELIEFS ARE NOT FUNCTIONAL STATES

Perhaps a functionalist could escape the contradiction as follows: Beliefs individuated by narrow semantic type must go the way of the indexical beliefs and de re beliefs already excluded from consideration; none is narrow enough to be a functional state. To put it another way, individuation by narrow semantic type (i.e., individuation by obliquely-occurring expressions in 'that'-clauses) may fail to coincide with individuation by narrow content (i.e., individuation without presupposing the existence of anything other than the subject).

The argument may proceed from referential semantics. The claim would be that states individuated by narrow semantic type are not narrow states, on the grounds that states individuated by narrow semantic type presuppose the existence of things other than the subject to whom they are ascribed. Individuation by narrow semantic type, according to the referential semanticist, is no narrower than individuation by truth-condition. But differences in truth-condition are not narrow differences; they are not "in the head", but rather are differences "in the world". Familiar arguments conclude that individuation by truth-condition does not coincide with individuation by functional role.14

Therefore, the argument goes, since differences in narrow semantic type
are exhausted by differences in truth-condition, it is no surprise that there can be differences in narrow semantic type without differences in functional states; again, this is so, it may be claimed, because functional states are narrower than the narrowest states individuated by narrow semantic type. Thus, the referential semanticist may cheerfully deny that he ever intended to identify functional states with narrow semantic types.

Without taking issue with this line of reasoning, let me just point out its implications for classical functionalism. In a word, if the functionalist claims that even beliefs individuated by narrow semantic type are too "wide" to be functional states, there seem to be no beliefs left to be candidates for functional states. Beliefs are taken to be attributable by 'that'-clauses; but if, as the referential semanticist urges, items individuated by the narrowest construal of 'that'-clauses (i.e., by narrow semantic type) are still too wide to be functional states, then no beliefs are functional states.\(^{15}\)

Therefore, individuation by 'that'-clauses fails to coincide with individuation by functional role. So, if, as functionalists hold, it is functional states that explain behavior (under the preferred descriptions), then beliefs and other states classified by 'that'-clauses do not. In that case, the entire practice of intentional explanation is jeopardized.

The conflict between individuation in terms of functional role and individuation in terms of 'that'-clauses seems insuperable.\(^{16}\) The discussion here has shown that the notion of functional state and the notion of belief state (classified by narrow semantic type) come apart. Thus, on the second horn, no belief that \(p\) is ever narrow enough to be an explanatory psychological state in the sense required by functionalism. Since on functionalism, (brute) behavior is to be explained in part by tokens of beliefs, to take this line is to abandon functionalism.

Although I have formulated the argument to reveal a conflict between individuation by 'that'-clauses and individuation by functional role, the general point that beliefs are not functional states may be made without recourse to individuation by 'that'-clauses.\(^{17}\) For the same point may be reached from another direction. Recall that a single functional state, at least in a computer, may equally well be a belief that cruise missiles can repel invasions and that pawns can ward off attacks, and notice that no appeal to individuation by 'that'-clauses is needed to recognize that the beliefs in question are different beliefs. So, without invoking individuation by 'that'-clauses, we again have the conclusion: beliefs can not be functional states.
To suppose otherwise would be to identify distinct beliefs.

To sum up the main line of argument: On the first horn of the dilemma, the functionalist is committed to an inconsistent triad. On the second horn of the dilemma, the functionalist argues that differences in narrow semantic type are really differences in truth-condition, and hence that belief states are not genuinely narrow states in the relevant sense. The latter approach has the immediate consequence that, since two tokens of a single functional state may be different (narrow) beliefs, beliefs are not functional states. Therefore, either beliefs are not psychological states, or psychological states fail to be functional states. In neither case is the functionalist’s optimism borne out.¹⁸

NOTES

¹ Following the tradition, I shall take belief to be the paradigmatic attitude. But what I say about belief clearly applies to any other “contentful” attitude ascribed by ‘that’-clauses. Intentions, desires, hopes, suspicions, fears, wishes, suppositions, thoughts, and so on, raise the same issues.

² Although the classical versions of functionalism construe functional states narrowly, there have recently been moves toward a “wide” functionalism, which individuates functional roles by reference to objects in the environment. See, for example, Patricia Kitcher (forthcoming). Also see Robert van Gulick (1982). For criticisms that apply to wide functionalism, see Joseph Owens, (1983). Although I believe that my arguments may be extended to wide functionalism, my concern here is functionalism as classically understood. On that construal, functional states are like machine states, individuated without reference to actual features of the external environment. (Consider SHRDLU, the famous program that simulates a block world, “manipulates” simulated blocks and pyramids, and answers questions about what it is “doing”. See Terry Winograd, 1972.)

³ Hilary Putnam (1975), p. 220. For a different approach to some of the issues I raise here, see Noonan (1981) and Morris (1984). By denying the intuition that explanatory psychological states must have semantic content, Morris, in effect, rules out attribution of beliefs as explanatory.

⁴ Fodor (1980), pp. 66–67. Hereafter, ‘MS’ will stand for the name of the article, references to which will appear in the text.

⁵ In Baker (1985, forthcoming) there is an argument for the following claim: if the narrow states that are to explain behavior are characterized as what molecular replicas share, then beliefs can not be explanatory. The conclusion here, based on different considerations, complements the conclusion of that paper. For another approach that concludes that beliefs can not be individuated “solipsistically”, see Garfield (1983).

⁶ There are reasons independent of functionalist commitments to suppose that such wide beliefs do not explain behavior. See, e.g., Perry (1979), Lycan (1981), Baker (1982a).

⁷ To see the nonequivalence, consider the theory according to which names are rigid designators with no semantic role other than to pick out objects. On that theory, either (i) the opacity of ‘S believes that Cicero was an orator’ and ‘S believes that Tully was an orator’ must be denied altogether, or (ii) beliefs that are different on (I) fail to be different on (II). (For a discussion of the first alternative, see Baker (1982b).) To preserve both opacity and the claim that beliefs different on (I) are also different on (II), we
can modify (I) by replacing the expression ‘co-extensive terms’ by ‘co-extensive nonrigid designators’. Since nothing I say here depends upon any particular semantic theory, I would accept the alternative formulation, noting only that to restrict (I) in this way is to undercut the motivation for (I) in the first place. These points are due to Alan Berger.

8 Lycan (1984a). A different version of these points may be found in Chapter Ten of Lycan (1984b), see also Lycan (1981).

9 Also see Jerry Fodor (1982).

10 Fodor (1981), p. 207. Although Fodor uses the illustration to criticize procedural semantics, the point is independent of any particular semantic theory.

11 Rey (1980).

12 This formulation of the proof was suggested both by Alan Berger and by Michele LaRusch.

13 The points in this paragraph emerged from a conversation with Jonathan Malino.


15 This argument from individuation is an argument against classical functionalism in principle. Some philosophers, such as Stich (1983), pp. 228ff, and Paul Churchland in (1979), p. 113, and in (1984), pp. 43ff., have been dubious of functionalism on more or less empirical grounds; they have entertained the possibility that concepts in the best scientific theory of the mind will not “match up” with intuitive concepts of belief. My arguments may be taken to show that if the best scientific theory of the mind construes explanatory psychological states as classical functional states, then there is a reason in principle for failure of one-to-one correspondence between the taxonomies.

16 In Fodor (1984), Fodor proposes a return to an observation/inference distinction. It may be possible to use such a distinction to argue that there are beliefs based directly on observation which may be individuated in a way that would qualify them to be functional states. I do not know how such a proposal would be worked out; and in any case, few of the beliefs attributed in intentional explanations are based directly on observation. See Baker (1985, forthcoming), Note 34.

17 This point was made by a referee for Philosophical Studies, who also noted that individuation of belief via ‘that’-clauses is a matter of controversy, despite its endorsement by functionalists.

18 For comments and criticism on earlier drafts of this paper, I am indebted to Alan Berger, William G. Lycan, Jonathan Malino, Igal Kvart, David F. Austin, Stanley Bates and Victor Nuovo. I am also grateful to the National Endowment for the Humanities and to the National Humanities Center for support for this project.

BIBLIOGRAPHY


———: 1985 (forthcoming), ‘Just what do we have in mind?’, in Midwest Studies in Philosophy 10, Peter A. French, et al. (eds.) (University of Minnesota Press, Minneapolis).


———: 1981, ‘Tom Swift and his procedural grandmother’, in Representations: Philoso-
A FAREWELL TO FUNCTIONALISM


1984a, ‘Semantics and methodological solipsism’, read at the Davidson Conference at Rutgers University.

1984b, Logical Form in Natural Language, (Bradford/MIT, Cambridge, MA).


Department of Philosophy,
Middlebury College,
Middlebury, VT 05753,
U.S.A.