



**WILEY-
BLACKWELL**



All The Difference in the World

Author(s): Tim Crane

Source: *The Philosophical Quarterly*, Vol. 41, No. 162 (Jan., 1991), pp. 1-25

Published by: Blackwell Publishing for The Philosophical Quarterly

Stable URL: <http://www.jstor.org/stable/2219783>

Accessed: 04/12/2009 06:55

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=black>.

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.



Blackwell Publishing and *The Philosophical Quarterly* are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Quarterly*.

<http://www.jstor.org>

The Philosophical Quarterly Vol. 41 No. 162
ISSN 0031-8094 \$2.00

The Philosophical Quarterly

Vol. 41 No. 162

January 1991

ALL THE DIFFERENCE IN THE WORLD

BY TIM CRANE

1. INTRODUCTION

The celebrated 'Twin Earth' arguments of Hilary Putnam (1975) and Tyler Burge (1979) aim to establish that some intentional states logically depend on facts external to the subjects of those states. Ascriptions of states of these kinds to a thinker entail that the thinker's environment is a certain way. It is not possible that the thinker could be in those very intentional states unless the environment is that way.

Those who accept this result are called variously 'broad minded', 'externalists', or 'anti-individualists'; those who reject it 'narrow minded',

‘internalists’, ‘individualists’ or ‘methodological solipsists’. There are many subtle differences in the positions that can be taken on this issue,¹ but there is a general consensus on the significance of the Putnam/Burge arguments. The consensus is that these arguments do succeed in forcing us one way or the other: for broad mindedness or against.

My purpose in this paper is to dispute this consensus. The Putnam/Burge arguments do not, I think, force us to opt for broad or narrow mindedness. I will argue not only that their conclusions are fundamentally opposed to crucial assumptions we are obliged to make about causation and the causal role of mental states, but also that the arguments for these conclusions are unsound. There is no Twin Earth problem of the kind Putnam, Burge and many others think there is. So there is no need to respond to it with broad- or narrow-mindedness.

A slightly different form of broadmindedness has been articulated and defended by Gareth Evans (1982) and John McDowell (1984, 1986). This position does not use Putnam/Burge-style Twin Earth examples to generate its conclusions, and for this reason I will not discuss it here. My concern in this paper is to refute the Putnam/Burge *arguments* for broadmindedness, not broad mindedness in every form.²

Why another paper on Twin Earth? My approach is, I think, different from that of most writers in the field. Most writers have accepted the Twin Earth arguments as sound, and have either developed theories of the mind to account for their conclusions, or altered their theories to accommodate them. But here I shall return to Putnam’s and Burge’s original arguments, and dispute them. My excuse for adding another paper to the already vast literature is that if my arguments are right, they will help not just to solve, but to *dissolve* the Twin Earth problem.

II. PUTNAM’S THOUGHT EXPERIMENT

Putnam’s original aim in ‘The Meaning of “Meaning” ’ (1975) was to dispute certain ‘grotesquely mistaken’ views of language (1975, p. 271) which arise from philosophers’ tendency to ignore the contribution made by our natural and social environment to the meanings of our words. Putnam claimed that these views depend on two incompatible assumptions about meaning. The first assumption is that knowing the meaning of a term is a

¹ In addition to Putnam’s (1988) and Burge’s (Burge 1982, 1986a, 1986b, 1986c) other work, three collections of papers stand out: Woodfield 1982, Butterfield 1986b (both largely narrow minded) and McDowell and Pettit 1986 (the broad minded backlash). For elaboration of some of the varieties of broad and narrow mindedness, see Blackburn 1984, ch.9, Butterfield 1986 and McGinn 1989, ch.1, esp. pp. 3ff.

² But see Segal 1989a for criticism of McDowell. For criticism of Evans, see Blackburn 1984, ch.9, Noonan 1986, and Carruthers 1987.

matter of being in a certain psychological state – in general, the meanings of words are fixed by the psychological states of those who use them. I shall call this ‘MPS’. The second is that meaning determines reference: a difference in reference is sufficient for a difference in meaning. I shall call this ‘MDR’.

Putnam uses the Twin Earth story to show that these two assumptions can be true of no notion – of meaning, Fregean *Sinn*, Carnapian intension or whatever. (I apologize to the reader for repeating the details of this familiar story, but it is necessary, I think, in order to disentangle some of the issues involved.) Putnam asks us to suppose that ‘somewhere in the galaxy’ there is a planet, Twin Earth, as similar as can be to Earth, except that on Twin Earth, the substance people call ‘water’ is not made up of H_2O , but has a complex chemical constitution whose description we may abbreviate to ‘XYZ’. XYZ feels and tastes like H_2O , and the people on Twin Earth do the same things with it. He also supposes that each of us has a duplicate ‘twin’ or *doppelgänger* on Twin Earth, type-identical to each of us down to the last atom.

Now suppose that on Earth I say ‘Water, water everywhere, nor any drop to drink’ and my Twin makes the same noises. Do we utter two sentences with the same meaning? Putnam argues that we do not, since the references of our words are different: H_2O and XYZ on Earth and Twin Earth respectively. Since the reference of the two utterances of ‘water’ is different on each planet, then by MDR, their meanings differ. But what each speaker ‘has in mind’ (sensations, beliefs about the superficial properties of water, etc.) is the same.

Putnam insists that this difference in meaning between Earth and Twin Earth does not depend on the fact that some scientists on each planet could *tell* that H_2O is not XYZ. To illustrate this he describes Earth and Twin Earth in 1750, before the development of adequate chemistry. In this case no one could tell the difference between the two substances; but the reference of ‘water’ on each planet differs, according to Putnam. And since the reference differs, so does the meaning, ‘in the intuitive, preanalytic use of that term’ (1975, p. 224). This is so on the plausible assumption that the meaning of ‘water’ does not change between 1750 and (say) 1950, simply because scientists found out more about water.

Putnam concludes that MPS is false. The psychological states of twins do not determine the reference of their utterances of ‘water’. So if we keep MDR, as he urges we should, then certain meanings aren’t determined by psychological states. Meanings aren’t ‘in the head’ (1975, p. 223).³

³ The ‘in the head’ locution has come in for a lot of criticism. As Davidson (1987) points out: to say that a certain condition of my skin is sunburn is to identify it by its cause. A twin with a type-identical skin condition which was not caused by the sun would not be *contd* . . .

The argument only establishes that MPS is false if my Twin and I are in the same psychological states. So the nature of what is being referred to (H₂O or XYZ) should not affect our psychological states in this sense. These are what Putnam calls psychological states ‘in the narrow sense’ (1975, p. 221), states which are permitted by ‘the assumption of methodological solipsism’:

that no psychological state, properly so called, presupposes the existence of any individual other than the subject to whom these states are ascribed. (1975, p. 220)

Psychological states which do not meet this condition are psychological states in the ‘wide’ or ‘broad’ sense: my Twin and I share our narrow psychological states, but differ in our broad psychological states. So what MPS really says is that the *narrow* psychological states of a language-user do not determine meaning.⁴

It is now obvious that there is an extension of Putnam’s argument to intentional states in general (Colin McGinn (1977) was the first to point this out). Assume that the contents of intentional states have truth conditions. Then if the truth conditions of my Twin’s and my states differ, so do their contents; and so do the intentional states we are in. (This is what it means to say that intentional states are ‘individuated by their contents’.) And this can happen, it is urged, without any change in our narrow states: so Putnam’s argument seems to show that intentional states, like meanings, are broad.

III. WHAT IS THE PROBLEM?

Why did Putnam’s argument create such a stir? He claimed that we should jettison MPS, but why did anyone believe it in the first place? After all, most philosophers of language were aware that reference – and hence meaning – had something to do with the relations between language-users

³ *contd.* sunburned. But this fact does not prevent my sunburn from being ‘in’ me. So it does not follow from the fact that a thinker’s state (e.g., knowing the meaning of ‘water’) is identified in terms of its causal history that the state is not properly a state of the thinker – not ‘in’ that thinker’s head. For this reason, I will refrain from talking about states being in or outside the head. There are clearer ways to put the broad/narrow distinction.

⁴ In order for Putnam’s definition of methodological solipsism to fit his story, narrow psychological states should fail to presuppose not just the existence of particular objects, but facts about the (e.g.,) chemical makeup of the substances they are about. For more on the definition of methodological solipsism, see Burge 1986b, pp. 112–14; Noonan 1984; Butterfield 1986 pp. 99–100. However, Twin Earth arguments can be used to support the thesis that some thoughts entail the existence of particular *objects* (see e.g., McCulloch 1986, pp. 62–3) but I think it is important to keep this thesis distinct from the precise claims of Putnam 1975.

and their environment, 'it takes two to make a reference' as David Lewis (1983 p. 371) has put it. So, since Putnam has done them the service of picking out some of these relations, why should anyone object to his conclusion?

One answer, which I do not endorse, comes from a 'Cartesian' conviction that narrow psychological states are the only ones that matter.⁵ The relevant Cartesian thesis is not substance dualism, but the idea that our states of mind are the way they are regardless of what the external world is like. How things seem to us is something to which we have immediate and authoritative 'access', and can be the way it is whatever the external facts are. This thesis can then be supported by some form of physicalism: being in an intentional state is a matter of being in a certain brain state, and one's brain could exist no matter what else exists. So, since I could have all the thoughts I now have in a world in which my brain is the only object, how could I have any mental states that entail the existence of something outside my brain? Thus Searle:

the brain is all we have for the purpose of representing the world to ourselves and everything we can use must be inside the brain. Each of our beliefs must be possible for a being who is a brain in a vat because each of us is precisely a brain in a vat; the vat is a skull and the 'messages' coming in by way of impacts on the nervous system. (1983, p. 230)

But this is precisely the picture Putnam is urging us to abandon. So it is no good to bring Cartesian 'intuitions' to bear in saying why the Twin Earth case is so problematic. For it is notorious that such 'intuitions' are often either deeply theory laden or as controversial as the thesis they are supposed to support. And since this Cartesian thesis is indeed extremely controversial, it will be best not to assume it in saying what is wrong with Twin Earth.

The problem with the Twin Earth result arises not from Cartesianism, nor indeed from physicalism, but from its conflict with two fundamental principles about causation and the causal nature of intentional states. The first principle is that intentional states have causes and effects. I take it as uncontroversial that beliefs, for example, are caused by perceptions and other beliefs, and combinations of beliefs and desires cause actions. This principle is crucial to the naturalistic aspirations of contemporary

⁵ See for instance Bach 1982, p. 127, and McGinn 1989 p. 33. Many broad minded philosophers have brought the charge of Cartesianism against narrow theorists – see, in particular, McDowell 1986.

philosophy of mind. Naturalism is the view that mental states are as much part of the natural world as physical, chemical, biological and other natural states. There is nothing about the mental that *prohibits* it from being described and explained in principle by a science of the mental, a science that invokes laws in much the same way that other sciences do. And if the natural is in some sense the causal, we have to believe that a naturalistic theory of the mental will, in part, be a theory of how mental states have causes and effects, and under which laws they fall.

It is worth saying that naturalism is not *physicalism*, the doctrine that ‘chemical facts, biological facts, psychological facts and semantical facts are all explicable (in principle) in terms of physical facts’ (Field 1972, p. 357). To think that mental facts are part of the natural world, and that they participate in causal interactions which may be explained by laws, in no way entails that they must be reducible to or ‘explicable in terms of’ physical facts. One can be a naturalist, and still believe that theories can define their own laws and concepts in their own terms, without them having to obtain a seal of approval from physics (see Crane and Mellor 1990).

There is of course much opposition to naturalistic theories of the mind. But I shall not argue for naturalism here, since my aim here is to address those philosophers who are naturalistic, but who also see Twin Earth as a problem to which they have to respond. I shall argue that, properly understood, Twin Earth provides no threat to naturalism; so naturalists can rest easy. However, my argument against Putnam will be of interest to non-naturalists too, if they think that Twin Earth shows that naturalism is doomed to failure (as Putnam himself does; see Putnam 1988, ch. 1).

The second principle is about which properties of things are involved in causal interactions between them. Suppose we assume, following Lewis (1983) and many others, that a property is the semantic value of a predicate. We need to make a distinction between two kinds of properties, thus liberally conceived.⁶ We should distinguish between those properties whose acquisition or loss by a particular is a *real change* in that particular, and those whose acquisition or loss is not. When I grow an inch and become six feet tall, there is a change in me. But when my brother becomes shorter than me by my growth, there is no change in him. His becoming shorter than me is a mere ‘Cambridge change’ (see Geach 1969, p. 71, and Shoemaker 1984, pp. 207–8). Properties of the first kind are, uncontroversially, *intrinsic* properties; and properties of the second kind are non-intrinsic.

⁶ Some writers have a causal test for real properties – e.g., Mellor 1981, and Shoemaker 1984 – and so would rule out being taller than my brother as a real property on these grounds. I am sympathetic to this approach, but it suits my expository purposes to begin by treating properties as the semantic values of predicates.

This distinction is important because when we look for causes, we look for intrinsic properties. (Whether these are properties of particular events or of particular objects is immaterial here). My brother's becoming smaller than me when I grew was not *caused* by my growth, and neither can the fact that my brother is smaller than me have any effects. But my becoming six feet tall was caused, and it can have effects: it can, for example, cause me to knock my head on a small door frame. My height, one of my intrinsic properties, is one of my *causally efficacious* properties; while my brother's being smaller than me, one of his non-intrinsic properties, is not one of his.

I admit that the intrinsic/non-intrinsic distinction is a problematic one, and I do not claim to define it in detail here. This is because all I need is the claim that there are clear cases of intrinsic properties and clear cases of non-intrinsic properties, and that a particular's causally efficacious properties are its intrinsic properties. The second principle relevant to Twin Earth, then, is that causation generally holds between instances of intrinsic properties.

Now the worry about Twin Earth is that broad intentional states are not, and do not systematically depend on, intrinsic properties of thinkers. Being in a world that contains H₂O is plainly not an intrinsic property of mine, and nor is being in a world that contains XYZ an intrinsic property of my Twin's.⁷ Yet according to Putnam, this difference *alone* gives our relevant intentional states different contents. What we are asked to believe is that my Twin's and my intentional states are different in virtue of a difference that does not depend on our intrinsic properties.

This would not matter if intentional states were not supposed to have causes and effects. There is nothing wrong with properties like being taller than my brother, or living in a world that contains XYZ. It is just that they are not causally efficacious properties. But believing that water is thirst-quenching should be, according to our naturalistic principle, a causally efficacious property. And on Putnam's account, the differences in the molecular constitution of water on Earth and Twin Earth make it the case that my Twin and I are caused to do different things – he drinks XYZ, I drink H₂O. In virtue of inhabiting worlds containing different watery stuff, our thoughts have different causal powers. But, if causally efficacious properties must be intrinsic, we must ask: how can there be such a difference in the causal powers of our states of mind without any mediating difference in our intrinsic properties?

⁷ We now come up against an irritating kink in Putnam's original story: our intrinsic properties *are* different, because most of our bodies are composed of water! I will follow most writers in treating this fact as irrelevant (see Dennett 1982, p. 11; McDowell and Pettit 1986, p. 2) so I hope the reader will bear with me in what follows.

This line of thought is implicit in Chapter Two of Fodor's *Psycho-semantic* (1987). Unlike Putnam, Fodor wants to defend methodological solipsism – though he uses the term in a slightly different way from Putnam. Fodor distinguishes between Methodological *Solipsism* and Methodological *Individualism* (1987, p. 42). The latter is, for Fodor, the empirical principle that sciences individuate kinds of entities by their causal powers. The former is the psychological claim that the contents of intentional states supervene on 'formal' properties of representations, items in a hypothesized 'language of thought'.⁸ Methodological Solipsism, on this construal, is equivalent to Fodor's 'formality condition' (1980, p. 227), the claim that mental processes only have 'access' to formal properties of mental representations: no mental difference without a 'formal' difference.

The route from Fodor's definition of methodological solipsism to Putnam's is quite straightforward if you assume, as Fodor does, that formal properties of representations supervene on intrinsic properties of the brains of thinkers. Since *ex hypothesi*, my Twin and I do not differ in our intrinsic properties, we do not differ in the formal properties of our representations either. And since we do not differ in these properties, methodological solipsism entails that our intentional states do not differ. My Twin and I therefore share all our intentional states if we share all our intrinsic properties: so intentional states must be narrow.

In my opinion, however, Fodor weakens his case by making the dependence on intrinsic properties a high-level generalization of scientific practice. This is a weakness because it is open for the broad theorist to respond that since the science of content is as yet waiting in the wings, we should not assume that it will not individuate content broadly (this seems to be Burge's (1986a) view). This response can be blocked by showing how the assumptions that Twin Earth challenges are not just generalizations about the current practice of science, but assumptions about causation and the causal status of states of mind.

This is not of course a knock-down argument against broad mindedness. For the whole point of the broad minded argument is that it is supposed to force a *revision* in our concept of an intentional state. What Twin Earth is supposed to show is that our intentional states depend on our environment in a much more radical way than was previously thought. So obviously some of our previous assumptions are to be challenged.

But my exposition of the problem makes it clear exactly what broad mindedness must deny in holding this. It must deny either that intentional

⁸ For the idea of a language of thought, see Fodor 1975, 1981, 1987 appendix; and Field 1978. I defend the coherence of the language of thought hypothesis in Crane 1990.

states as such have causes and effects, or that relational properties like inhabiting a world in which water is XYZ cannot have immediate effects.⁹ That is, to be broad-minded, you must give up one or both of the principles I outlined above. But both principles are, I think, extremely compelling. So if the Twin Earth argument were sound, it would undermine much of what we think about causation and the causal status of mental states. This is why it matters.

IV. WHAT IS WRONG WITH PUTNAM'S ARGUMENT?

I now turn to the assessment of Putnam's argument. Despite its influence, I think that reflection shows that it is inadequate, and that his positive claims are either implausible or largely irrelevant to the problems of intentionality and mental representation. However, to be fair to Putnam, I am only interested in his argument in so far as it bears on intentional states and their contents. I think some of the points he makes about language and its social character are well taken, but part of what I shall argue is that we should not be hasty in reading off claims about intentional states from claims about the use of language.

Putnam's argument rests on the following three premises:

- (1) MDR – meaning determines reference.
- (2) My Twin and I are atom-for-atom identical, and so (by definition) share all our narrow states.
- (3) When I say 'water is wet' and my Twin says the same, our sentences do not have the same truth conditions, since his water is XYZ and mine is H₂O.

From these he concludes that narrow states – the states my Twin and I share – do not determine the meanings of our words.

The argument is valid, so we must look at the premises. Premise (1) is just the assumption MDR, which we should accept (*pace*, perhaps, the case of indexicals¹⁰). Premise (2) is part of the story, and although it is rather

⁹ It is clear what Putnam thinks: 'Only if we assume that psychological states in the narrow sense have a significant degree of causal closure . . . is there any point in . . . making the assumption of Methodological Solipsism. But three centuries of failure of mentalistic psychology is tremendous evidence against this procedure, in my opinion' (1975, p. 221). Putnam 1988 develops this line of thought in a critique of functionalism.

¹⁰ The idea being, of course, that for an indexical, meaning-plus-context determines reference (see Kaplan 1977, Perry 1979). See below on Putnam's application of the semantics of indexicals to the 'water' case. Fodor has attempted to generalize this feature of indexicals to all contents to account for Twin Earth: 'narrow' content, he says, is a function from context to truth-conditions (1987, ch. 2). The difficulties for this proposal are many – but if my approach to Twin Earth is taken, there will be no need to introduce the notion of narrow content, as we shall see.

irritatingly undermined by the fact that our bodies are composed of water, this is just a feature of the example. There is certainly nothing incoherent in supposing that there can be two people atom-for-atom indistinguishable.

So all the interest rests with (3). Why should we believe it? Why should a difference in the chemical structure of water affect the truth conditions of 'water is wet'? Why should we not say, for instance, that if there were such a substance as XYZ, all this would show is that 'not all water has the same microstructure' (Mellor 1977, p. 303). And indeed, we knew this already, since it would surely be stipulative to deny that heavy water (D_2O) is really *water*. So why does the fact that most of our water is H_2O entail that the truth conditions of 'water is wet' differ across Earth and Twin Earth?

Perhaps a defender of Putnam could respond that 'water' is not really a natural kind term. Perhaps we should accept that there are other kinds of water, because of the different chemical structures that can have similar superficial properties. But when you get to the names of elements you get 'real' natural kind terms – there cannot be different kinds of gold, or lead or helium.

But what about isotopes? Which isotopes of elements are the substances referred to by the 'real' natural kind terms? Which of the two isotopes of chlorine is really chlorine? Maybe they both are – but in that case 'chlorine' like 'water' is not a real natural kind term either, for the same reasons. And this seems to restrict the range of real natural terms to such an extent as to make them a trivial category for metaphysics and the philosophy of science: terms that pick out elements that do not have isotopes. The obvious lesson is that the idea of a natural kind term, as used by essentialists like Putnam, is not very well defined.¹¹

It may be responded that when there *is* a clear case of two genuinely different elements – not isotopes of the same element – that are superficially indistinguishable, then Putnam's argument will work. This does somewhat reduce the force of the argument, but it is worth looking at such a case in some detail to see exactly what is wrong with it.

Take aluminium and molybdenum, two practically indistinguishable metals whose names (and relative scarcity) are switched on Earth and Twin Earth. And suppose my Twin and I are atom-for-atom identical. Putnam says that the meaning of 'aluminium' spoken by my Twin and me is underdetermined by our narrow states – the states we share. Our narrow states determine molybdenum *and* aluminium as the extension of our uses of the word 'aluminium'. But why not therefore say that neither of us has a full understanding of the meanings of our words? Why not say that we have

¹¹ For a refutation of the Kripke/Putnam view of natural kinds, see Mellor 1977. I am indebted to this paper, and to Dupré 1981.

the same concept (call it '*molyminium*') that applies to aluminium and molybdenum alike? The concept *molyminium* will distinguish less 'finely' between substances than the concepts *aluminium* and *molybdenum*.

That there can be incomplete understanding of the meanings of words in natural language is indisputable. Indeed, it must be a premise in Putnam's own argument for the 'division of linguistic labour' (1975, p. 227) which I will discuss briefly below. But Putnam may respond that my 'common concept' response forces us to give up MDR. Since the extensions of my Twin's and my use of 'aluminium' are different, then the common meaning of our utterances – the concept *molyminium* – does not determine the extensions of our uses.

I agree that the concept *molyminium* does not determine the exact extension of the Earthly word 'aluminium', and nor does it determine the exact extension of the Twin Earthly word. But it still determines *an* extension: the extension that includes both aluminium and molybdenum. Since my Twin and I have an incomplete understanding of the word 'aluminium', our concept *molyminium* does not determine the extension of 'aluminium' as the concept of someone who does have a complete understanding does. Such a person, someone 'in the know', has the more precise concept, and is thus in a position to correct the uses of the word 'aluminium' of those speakers who do not distinguish sufficiently between aluminium and molybdenum. This is just Putnam's division of linguistic labour – but it gives us no reason to think that my Twin and I do not share our concepts. Indeed, it gives us a reason for thinking that we must share them.

The reason for this is that we should distinguish between the conventionally assigned meaning of a word in a public language, and the concept intended to be expressed by the user of that word. This distinction should not be controversial – it is needed, for instance, to make sense of ambiguity and punning. The sentence 'The ship is veering to port' has at least two meanings in standard English, but a sincere competent assertion of it could only be the expression of one content (the example is from Evans 1982, p. 68). There can be unclarity in thought, but not ambiguity.

This point is actually very important if we believe, as I think we should, that thought can be independent of public language. For after all, incomplete understanding is still understanding – there is the world of difference between my utterance of 'Planes are made of aluminium', even when I cannot distinguish between aluminium and molybdenum, and the same sounds made by a monolingual Chinese speaker or a parrot. As we say, the Chinese speaker and the parrot have 'no idea what they are talking about'. But I do have an idea – and my Twin has the same idea – it is just not as precise an idea as that of someone in the know.

So my concept *molyminium* is not the same as the concept of someone in the know – even when the same word is used by him and me. When I say ‘Planes are made of aluminium’ on Earth, those in the know will take me as talking about aluminium. So I will have succeeded in referring to aluminium – after all, my concept *molyminium* does include aluminium in its extension. But since there is nothing in my mind that enables me to distinguish between aluminium and molybdenum, my concept includes molybdenum in its extension as well. Similarly with my Twin – except he succeeds in talking about molybdenum, since this is how those in the know on Twin Earth would take his words. So on a number of plausible assumptions, the ‘aluminium’ example does not show that my Twin’s and my intentional states differ purely in virtue of the difference in the relative scarcity of aluminium and molybdenum, and the difference in the use of a word. So it does not show that intentional states are broad.

But what if no one is in the know? Putnam’s claim about the difference in the meaning of ‘water’ on Earth and Twin Earth in 1750 is intended to show that this does not matter. The meaning of ‘water’ depends on its microstructure, and since H_2O is not XYZ, the meanings of utterances of ‘water’ vary across the planets no matter who is in the know.

This cannot be right. Even if we ignore the largely unsupported assumption that the microstructure of a substance gives the meaning of a word for that substance, the point falls foul of the principle, mentioned in §III, that immediate causation must go via intrinsic properties. Consider a ‘visiting case’ when a pre-1750 Earthian travels – without his knowledge – to Twin Earth. On Putnam’s view, his intentional states that are about water will be changed purely in virtue of the fact that he is now related to XYZ. But if our causal principle is right, then this cannot be a real change unless there is some change in his intrinsic properties, which *ex hypothesi* there is not. So Putnam must say that the change in his intentional states is not caused by the presence of XYZ – but there is surely little reason to accept this conclusion, given the availability of the ‘common concept’ strategy.¹²

¹² Colin McGinn has argued that Putnam’s argument depends on the ‘highly compelling and intuitive thesis’ that ‘perceptual seemings are not necessarily as fine-grained as the reality that causes them’ (1989, p. 32). For him, the argument is a special case of the plausible claim that how things seem to us does not determine how they are: for a thinker’s thoughts about the superficial properties of water (the ‘seemings’) do not alone determine the reference of those thoughts (the ‘reality’). But Putnam’s argument needs more than this. To hold that XYZ and H_2O are both water – and therefore to say that my Twin and I share our water-thoughts – is consistent with holding that the relevant seemings can be mistaken – my Twin and I can both be wrong about whether what we have before us is indeed water. To get to Putnam’s conclusion, we have to add the tendentious assumption, which I deny, that the relevant feature of the reality in question is the microstructure of water.

I conclude that Putnam's argument fails. If there are people in the know, then the public meaning of my words can 'float free' of my concept, as Putnam himself admits. But this does not prevent me and my Twin having the same concept. But if no one is in the know, the concept expressed on Earth and Twin Earth by 'water' applies to H_2O , XYZ, D_2O and anything else that is – how shall we put it? – *water*.

What about Putnam's positive proposals? The two major claims made about meaning are that natural kind terms like 'water' are indexical, and that there is a division of linguistic labour (a thesis I mentioned above). These two claims are intended to compensate for philosophers' lack of concern for the roles of 'the world' and 'other people' respectively in determining meaning. However, I think that the first claim is both false and inconsistent with other claims of Putnam's, while the second is true but irrelevant to this paper.

First, the Twin Earth cases are meant to demonstrate that the world itself can, as it were, fix the meanings of some of our words. To show how this might work, Putnam introduced the idea that natural kind words have an 'unnoticed indexical component: "water" is stuff that bears a certain similarity relation to the water round here' (1975, p.234). The idea seems to be that once the reference of 'water' has been fixed (perhaps by saying 'This is water'; (Putnam 1975, p.231)) then the word can only refer to the substance which is relevantly similar in constitution to the substance referred to. So if meaning still determines reference, then the meaning of a term must be sensitive to the context of the use of that term. 'Context' is often a vague term in this debate – a catch-all label for those things we do not understand – but in the 'water' case, it is fairly clear that for Putnam, the relevant features of the context are the microstructure of water, and where the utterance is made.

It is a commonplace that indexical terms fix reference by sensitivity to context. 'I', 'here' and 'now' refer to different speakers, places and times in different contexts of utterance. This is clearly what makes Putnam say that natural kind terms are indexical. And he also sees a connection between this idea and Kripke's thesis that natural kind terms are 'rigid designators': they refer to the same substance in all possible worlds in which that substance exists (see Kripke 1980 pp. 3–5; Putnam says that Kripke's doctrine and his are 'two ways of making the same point'; 1975, p. 234). So, since the Twin Earth cases are supposed to show that there is no water on Twin Earth, Putnam concludes that 'water' works something like 'I' in picking out different substances in relevantly different contexts.

However, even if we accept that XYZ is not water, this claim is incorrect. For if 'water' is a rigid designator, then 'water' as used on Twin Earth is a mere homonym of 'water' as used on Earth. But indexicals are not

homonyms – indexical type-expressions have a constant meaning. For example, since the work of Kaplan (1977) and Perry (1979), many accept that the constant meaning of (say) the type ‘here’ is a function which maps contexts (places of utterance) on to a contribution to truth conditions.¹³ But Putnam should hold that there is *no* constant meaning in the uses of ‘water’ across Earth and Twin Earth, since he believes that meaning determines extension. (This is a point made by Burge (1982, p. 102).)

So if ‘water’ is genuinely indexical, then all tokens of ‘water’ must have a constant meaning. But this is inconsistent with the conclusion of the Twin Earth argument. On the other hand, if the Twin Earth argument is sound, then tokens of ‘water’ spoken on both planets are mere homonyms. But this is inconsistent with the claim that ‘water’ is an indexical. So if Putnam is to keep the conclusion of the Twin Earth argument, then he ought to abandon the claim that ‘water’ is indexical.¹⁴

Putnam’s second positive claim is based on his suspicion that the *communal* aspect of meaning has been ignored by philosophers of language. As I said above, it is plain that for many of the words we use, we would be unable to say exactly how to apply them. Scientific words which have passed into everyday usage, like ‘protein’, ‘gene’, or ‘evolution’ would be good examples. According to Putnam, our ability to use these terms depends on a ‘division of linguistic labour’ (1975, p. 227). The users of certain terms – Putnam talks only of natural kind terms – are divided into ‘producers’ and ‘consumers’: there are those who have a precise grasp on how to apply a particular term, and there are those who use the term without having such a grasp. In using such terms, the consumers ‘defer to the experts’, who have authority on what the meaning of the term is.

I think this is right, but it does not tell us much about intentional states. The division of labour is a consequence of the communal nature of language and inquiry. Consider the Quine–Neurath metaphor of the community of enquirers as sailors on a ship: some of us replace the planks in the hull, some of us weave new sails, some of us navigate, some of us repaint the figurehead and some of us cook. No one sailor can do everything on such a large ship and still keep it seaworthy – hence the division of epistemic and linguistic labour. If you are mending the sails, wait for the cook to provide your meals. But don’t waste your time wondering what that fish soup is made of – or what the words ‘Clam Chowder’ really *mean* – we have a ship to keep afloat!

¹³ I am not denying that it is a difficult task to explain how a token indexical refers in a context – for some discussion, see Perry 1979, Kaplan 1977, Evans 1982 and Mellor 1989.

¹⁴ The fact that the reference of ‘water’ may be fixed by a ‘baptism’ *using* an indexical is irrelevant, since Putnam will surely agree with Kripke (1980, pp. 55–8) that fixing the reference of a term does not amount to giving its meaning.

This whole picture depends on the idea that there can be incomplete understanding of the meaning of words. But as I said above, where there is incomplete understanding, there is still *some* understanding – and it is this we are interested in when analysing intentionality. The meanings of words in a communal language can float free of the beliefs of individual speakers – but this uncontroversial description of the division of labour *presupposes* the very distinction between public meaning and belief that shows the thesis to be about public linguistic meaning, not about the contents of individuals' intentional states. In stating the thesis we have to *assume* that it is not simply about the contents of intentional states; this is why, though true, it is irrelevant to this paper.

So Putnam's main claims in 'The Meaning of "Meaning"' are either implausible or largely irrelevant to intentionality. In particular, they do not seem to show anything about the broadness of intentional states unless certain implausible assumptions (e.g. essentialism about natural kinds) are made. It has been worth spelling this out, since many writers, on both broad and narrow sides, simply assume that Putnam has shown that the 'ordinary notion' of an intentional state is a broad notion (e.g., Fodor 1980, Butterfield 1986, and Jackson and Pettit 1988). If I am right, the argument does not begin to set up the problem to which so many writers have felt the need to respond. However, perhaps Burge's version of Twin Earth fares better.

V. BURGE'S THOUGHT EXPERIMENT

Burge has used a version of Putnam's Twin Earth story to argue against 'individualism', the view that 'there is no necessary or deep individuating relation between the individual's being in [intentional] states . . . and the nature of the individual's physical or social environment' (Burge 1986a, p. 4). Burge is clearly using the term 'individualism' in a different way from Fodor (see §II above). For it may be that individualism in his sense is false while in Fodor's sense it is true, and vice versa. This would be so if the nature of an individual's physical or social environment had a 'necessary or deep' effect on the causal powers of a thinker's thoughts.

Burge seems not to believe that intentional states must be individuated by causal powers (1986a, p. 16), but this is not the same as denying individualism in his sense. The real issue between Burge and Fodor is rather over the truth of Fodor's methodological *solipsism*: the claim that intentional states supervene on formal (i.e., intrinsic) properties. For once this 'formality condition' is accepted, then differences in intentional states can only come about through differences in intrinsic properties. Burge's thought experiment is intended to show why this is false.

Here is Burge's familiar story.¹⁵ Suppose that Alf has a large number of beliefs and desires about arthritis. Many of Alf's beliefs about arthritis are true, but he also has the false belief that he has arthritis in his thigh. His belief is false because although he 'grasps' the concept of arthritis, he misapplies it in this particular case: arthritis can only occur in the joints. Now imagine a counterfactual situation in which all Alf's physical and non-intentional psychological states are the same, but in which the word 'arthritis' is applied by Alf's community to many different types of rheumatoid ailments, including the disease Alf has in his thigh. In this case, Alf's belief would be true. Burge's conclusion is that in the counterfactual situation, Alf lacks some – perhaps all – of the 'attitudes commonly attributed in the actual situation with content clauses containing the word "arthritis" in oblique occurrence' (1979, p. 75). In other words, in the counterfactual situation Alf has no beliefs (or desires, etc.) about arthritis. This is because he does not have the *concept* of arthritis in the counterfactual situation. He has a different concept, which he expresses by using the word 'arthritis'. Yet *ex hypothesi* Alf's intrinsic properties have not changed in both situations.

Burge's thought experiment differs from Putnam's in two significant ways. First, it is concerned directly with the contents of intentional states, not just with word-meaning; an extension which, as we saw, Putnam could have made himself. And second, it does not just apply to 'natural kind' terms, but it can be applied to any term that someone could mistakenly misuse (see Burge 1986c). The general point of the argument is that we can conceive of counterfactual situations in which all the physical and non-intentional psychological states of a thinker are the same as in an actual case, but in which the thinker's intentional states vary because of a difference in some feature of his or her linguistic environment.

VI. WHAT IS WRONG WITH BURGE'S ARGUMENT?

Burge's argument raises the same worry as Putnam's. If immediate causation acts only via intrinsic properties of things, how can a difference in the contents of two subjects' intentional states have any causal manifestation if there is no intrinsic difference between the subjects? Burge thinks that the intentional states of actual and counterfactual Alf are different, and that this difference is due simply to a difference in the uses of words in their linguistic communities. It has no reflection in their intrinsic

¹⁵ Burge has given many versions of his story, some more elaborate than others (1986c gives the most recent and elaborate version). For my purposes, these differences do not matter, since I want to capture the essentials of his argument.

properties. Of course, there are many links between thinkers' thoughts and the linguistic practice of their communities – and many of the links are surely causal. The problem for Burge is to say how this can happen without affecting thinkers' intrinsic properties.

In his most recent criticism of Burge, Fodor (1987) has a similar objection. He wants to defend mind/brain supervenience, which he thinks is 'the best idea that anyone has had so far about how mental causation is possible' (1987, p. 30). Burge (1986a) rejects supervenience, though he does believe that intentional states supervene on physical properties of the brain-plus-environment. Fodor objects, in effect, that Burge's thesis provides no way of accounting for the fact that causal relations hold between intrinsic properties of things:

you can't affect the causal powers of a person's mental state without affecting his physiology. That's not a conceptual claim or a metaphysical claim of course. It's a contingent fact about how God made the world. God made the world such that the mechanisms by which environmental variables affect organic [*sic.*] behaviours run via their effects on the organism's nervous system. Or so, at least, all the physiologists I know assure me. (1987a, pp. 39–40)

Burge will respond that Fodor's objection begs the question by simply assuming the truth of supervenience; after all, part of what Burge wants to establish is that there is a way in which thinkers' intentional states can be 'affected' without affecting their physiology. But as we saw in §III, Fodor's objection can be made more powerful, and more general, by recognizing that the point about the causally mediating properties is not so much that they should be physiological, but that they should be *intrinsic*. (And I suppose this is a 'conceptual or metaphysical claim'.)¹⁶

Burge responds to the Fodorean objection by saying that it confuses 'causation with individuation' (1986a, p. 16). Indeed it does associate causation with individuation; but there is no confusion. Fodor is absolutely right to say that theories do, and should, individuate properties by their causal powers or dispositions. There will always be other ways of describing and picking out properties, but when our concern is with what happens and why, we should look for causal powers.

Given this objection, it will be instructive to analyse the stages in Burge's actual argument, to see what is wrong with it. In doing so, we will learn

¹⁶ Putting the point this way isolates Fodor's objection from the commitment to mind-brain supervenience – which may make it more acceptable to some. For doubts about supervenience, see Crane and Mellor 1990, §5.

something about the relation between thought and language, just as we did when discussing Putnam.

As we saw, Burge's thought experiment is supposed to work using any concept that a thinker could mistakenly apply. To begin with, we should distinguish between what I shall call Burge's *linguistic* and his *non-linguistic* thought experiments. Although he emphasizes the role of the linguistic community in shaping content, Burge does not want to argue that one needs to have a language in order to have intentional states. Indeed, in his more recent papers, he has tried to formulate his thought experiment in such a way that it applies to all types of content, not just that expressed or expressible by language-users (1986a; 1986b). But for this purpose he has to introduce a different set of considerations, to play an analogous role to the one played by the practice of the linguistic community in the original thought experiment. So we should consider the linguistic and non-linguistic cases separately.

To see what is wrong with Burge's linguistic thought experiment, I suggest that we should look more closely at the source of the alleged error in actual Alf's belief. In order for Alf to express his belief with the sentence 'I have arthritis in my thigh', not only does he have to believe that the disease he has in his thigh is arthritis, but also that the sentence 'I have arthritis in my thigh' is the right one to express this belief. For beliefs to be expressed in words, they have to go via second order beliefs about which words are the right ones for expressing which beliefs: sentences do not, as it were, just 'squirt out' beliefs.

So Alf's error is introduced either by the belief (1) that *I have arthritis in my thigh*, or by his belief (2) that '*I have arthritis in my thigh*' is the right sentence to express this belief, or by both. (In (2), 'this belief' refers not to (1) but to whatever first-order belief Alf is having about the disease in his thigh.) Clearly, the latter belief (2) is false. But what about the former? To describe Alf's beliefs correctly, we have to choose between: first, attributing to him a concept *tharthritis* which applies to both arthritis and whatever is the disease he has in his thigh; and second, attributing to him the public concept *arthritis*. The first attribution would entail that his first-order belief is really *I have tharthritis in my thigh* and is true; the second would entail that it is belief (1) and false.¹⁷ Which attribution is right?

Linguistic evidence is not enough to decide this. For as we saw, the falsity of Alf's belief that '*I have arthritis in my thigh*' is the right sentence to express this belief is enough to explain why his utterance of 'I have arthritis in my thigh' is false. So the utterance's falsity cannot be used as evidence in

¹⁷ I claim no great originality for the 'tharthritis' move; it is a natural response to Burge, and one that he has not, I think, adequately refuted.

deciding between the first and the second attributions. What about non-linguistic evidence? It has been persuasively argued by Ned Block that a problem with Burge's thesis is presented by the fact that a psychologist would not be able to distinguish between the non-verbal behaviour (e.g., food preferences, in the 'brisket' example) of two Burgean twins in certain experimental situations (Block is referred to by Fodor 1987, p. 40). This would seem to be so in our case too: up to the time described by the thought experiment, Alf has (*ex hypothesi*) all the same dispositions to (non-verbal) behaviour in the actual and the counterfactual situations. This suggests that we should attribute the concept *tharthritis* to him.

And indeed this does seem to be the right thing to do. For until actual Alf is able to correct his belief about the meaning of the word 'arthritis', there is no reason as yet to suppose that he can discriminate between arthritis and tharthritis. So his concept – which at the very least must reflect an ability to discriminate – will apply to arthritis and tharthritis alike.

On this diagnosis, then, Alf has a true belief, *I have tharthritis in my thigh*, a false belief to the effect that '*I have arthritis in my thigh*' is the right sentence to express this belief, and thus makes a false statement, 'I have arthritis in my thigh'. This means that though his belief is true, he says something false when he attempts to express it. This sounds paradoxical, but it becomes clear when we distinguish, as we did when discussing Putnam, between the meanings of sentences in public languages and the contents of beliefs.

As we saw above, once words in a language are endowed with meaning, there 'opens up the possibility of a gap between what a speaker means to say by uttering certain words – what [content] he wishes to express – on the one hand, and what he strictly and literally says, according to the conventional meanings of the words he utters, on the other' (Evans 1982, p. 67). As we saw, the existence of ambiguity and punning clearly show the need for postulating such a gap. Also, I can say something by uttering 'p' without knowing the precise meaning of 'p' and thus without expressing a belief of mine by my utterance. It is therefore quite unsurprising that there should be cases, like Burge's, where a thinker may have a true belief, *p*, but yet utter a falsehood because of a mistaken belief about which words are the right ones to express *p*.

The key idea here is that the expression of beliefs in language is mediated by beliefs about which words to use. And this is a point which I think Burge does not sufficiently appreciate. In his paper, 'Belief and Synonymy', he argues that we can quite easily make sense of someone who says

For years I believed that a fortnight was ten days, not fourteen, though of course I never believed that fourteen days were ten days. (1978, p. 126)

without having to attribute him any ‘metalinguistic’ beliefs about the meanings of his words. Now indeed this remark makes perfectly good sense, and one who utters it does not say something contradictory. But surely this is *only* because we should construe the speaker’s previous erroneous belief as involving a mistaken belief about the meaning of ‘fortnight’. What the speaker believed was that ‘*fortnight*’ means a period of ten days. What else could his belief be? There is nothing more to being a fortnight than being a period of fourteen days; there would be nothing *about fortnights* for the speaker’s beliefs to latch on to, apart from the definition of ‘fortnight’.

Burge considers this response but dismisses it, chiefly on the grounds that (a) thinkers need have no beliefs about meanings in order to think a fortnight is ten days (see, e.g., 1978, p. 126); and (b) this way of construing the belief ‘does not accord with our ordinary attributions of belief’ (1978, p. 132). Neither point is persuasive. It is true that one needs no beliefs about meaning, truth or reference in order to think; but to express one’s thoughts in words, it is not possible that one should lack beliefs about these things. Imagine asking a speaker, ‘Do you believe that the word “fortnight” means a period of ten days?’ The speaker will answer ‘Yes’, ‘No’ or ‘I’m not sure’ – in each case, expressing a belief about the meaning of a word.

As for (b), this is simply debatable. Burge seems to treat the particular sentence used by a thinker to express a belief as a transparent medium, through which the ascriber of a belief can simply gaze and read off its content.¹⁸ But surely the business of ascribing beliefs is rarely like this. In working out what others think, we often need to question them further, translate their use of words into our ‘idiolects’, and redescribe their attitudes as evidence accumulates. We do not, as Burge urges we should (1978, p. 132), take utterances like the one above at face value, any more than we would take a speaker’s sincere utterance of ‘I have a hippopotamus in my refrigerator’ at face value (Davidson 1984, pp. 100–1).¹⁹

So this is why I think it plausible to say that in our *arthritis* example, Alf has one true belief, one false one and makes a false (public language) statement. Of course there are, as a matter of fact, concepts whose possession requires knowing a public language. But this does not yield Burge’s conclusion, since it is consistent with holding that the relations

¹⁸ Here I agree with Kent Bach, who challenges Burge’s ‘assumption that when we use a term in the “that”-clause of an attitude attribution literally and correctly, we must be ascribing the notion [concept] expressed by the term to the content of the attitude’ (Bach 1988, p. 88).

¹⁹ The point derives from Quine: ‘Assertions startlingly false on the face of them are likely to turn on hidden differences of language. The maxim is strong enough in all of us to swerve us even from the homophonic method that is so fundamental to the very acquisition and use of one’s mother tongue’ (1960, p. 59).

between a community's linguistic practice and an individual's behaviour have to be mediated by their intrinsic properties – a thesis Burge implicitly denies.

However, it appears that if this analysis of Burge's case is right, then thinkers cannot make certain sorts of mistake.²⁰ I said that it was more plausible to treat Alf and counterfactual Alf as having the same concept, *tharthritis*, with the result that they both have true beliefs about *tharthritis* while actual Alf has a false belief about what 'arthritis' means. But, it may be said, it seems that if this procedure is generalized, we can always explain away the appearance of an error in a thinker's belief, when that belief is not about the meaning of a word, by attributing to him a different concept. So when someone mistakenly thinks that *a is F*, we can explain away this error by attributing to him a non-standard concept *F**, which has *a* in its extension. No one should doubt that thinkers can have non-standard concepts; but it is equally undeniable that thinkers make simple mistakes. It might appear that my analysis of Burge's case assimilates the second sort of case to the first: there can be mistakes in perception, and in beliefs about what words mean, but there cannot be mistakes in the beliefs in between.

This would be an unacceptable result. We often make mistakes in thought, and a theory of thought must explain this. But in fact my analysis of Burge's case can cope with this, as long as the right distinctions are made. When working out what thinkers believe, we have to take into account not only the evidence of what they say or do at a particular time, but what they *would* say or do under other circumstances. That is, we have to consider which counterfactuals are true of them. Taking only the evidence of one utterance or one action will, of course, radically underdetermine the correct ascription of the thinker's beliefs.

Such limited evidence will also underdetermine whether a thinker makes a genuine mistake or has a non-standard concept. Suppose we have *prima facie* evidence that a thinker, T, believes that an object, *a*, has a certain property, *F*, when it does not. What settles whether T falsely believes that *a is F* or truly believes that *a is F** is just which counterfactuals are true of T. If T would apparently believe, when confronted with *a* on another occasion, that it is *F*, then we might be tempted to think that he has a concept *F**. If T would (apparently) believe, for instance, that some things that are similar to *a* are also *F* (when they are not), then this would support the attribution of *F**. If T is a sophisticated thinker, T may, under questioning, offer other beliefs of his about *a*, connected in more or less consistent ways. This would again support the attribution to T of the belief that *a is F**. We may be thereby entitled to attribute to T a slightly off-beat

²⁰ Here I am grateful to Mark Sainsbury.

theory, in which the concept F^* figures, and applies to a and things similar to it.

If, on the other hand, T would not be disposed to believe that some things similar to a are F , and would not have many other beliefs to call upon in support of his (apparent) belief that a is F , then T's belief that a is F is a mere 'slip of the mind' – a mistake, but one that does not infect the identity of the concept. It is hard, without a theory of concepts and of error, to say here exactly which counterfactuals are essential to having the belief that a is F at all. But what matters here is that if it is granted, as it should be, that the two sorts of case are different, then my analysis of Burge's thought experiment will go through. This is because Burge's 'arthritis' case is much more like the first case than the second. We are told that actual Alf has lots of true beliefs 'about arthritis'. He thinks he has a disease of the joints, that it is very painful, and that he has the same disease in his thigh. So he thinks the disease he has in his thigh can occur in muscles. He has an off-beat theory of the various ailments he has – not a very well-informed theory, but a theory none the less. This is why I say he has the concept *tharthritis*, one that he shares with his counterfactual 'twin'. Of course, since this concept does not fit the usage of actual Alf's community, it is the wrong concept for him to have: he is wrong to think that there is such a disease as *tharthritis*. But what I am urging is the importance of distinguishing this belief from the belief that he has arthritis in his thigh – a belief that neither Alf nor his counterfactual twin has.²¹

Much more needs to be said: about how, in general, error comes about, about how to distinguish, in a principled way, between errors and non-standard concepts, and about how to individuate concepts. None of these questions have easy answers, but all I need for the purposes of this paper is to show why my analysis of Burge's cases does not rule out the possibility of any error. And this much, I claim, can be done.

So that is how I deal with Burge's linguistic thought experiment. But Burge also wants to show that even intentional states which are not essentially expressible in a public language are non-individualistic. However, I shall not deal with Burge's non-linguistic thought experiment, and its ingenious application to Marr's theory of vision, since I think these arguments have been refuted by Gabriel Segal (1989b) and I refer the reader to the details of Segal's paper. My target in this paper has been Burge's linguistic thought experiment.

²¹ Again, my sympathies are with the letter, though not the spirit, of some of Davidson's remarks on interpretation. Davidson says that he does 'not find Burge's thought-experiment as persuasive as [Burge] does; they seem to me at best to encourage us to consider what principles we use in (correctly) interpreting the thoughts and words of others' (1988, p. 665).

VII. NEITHER BROAD NOR NARROW

So Burge and Putnam fail to provide good arguments for the conclusion that intentional states are broad. Nor do they show that the constitution of natural kinds or the linguistic practice of a community affects the identity of intentional states in the way their thought experiments say. But this does not mean that we should retreat from broad mindedness into a narrow minded position. In particular, we should not suppose that what my Twin and I (or Alf and counterfactual Alf) have in common is an attitude to a ‘*narrow content*’ (see Fodor 1987, ch. 2; Block 1987). Nor should we be thereby committed to a ‘Two Component’ theory of content which says that intentional states are composite states, involving a broad ‘component’ which, as Twin Earth shows, determines the reference and truth conditions of the state’s content, and a narrow, ‘internal’ component which is responsible for the causation of behaviour (see McGinn 1982, pp. 208–16).

The reason why narrow mindedness is not the alternative to the Putnam/Burge thesis is that narrow mindedness, and the notion of narrow content, were only introduced in response to the challenge of Twin Earth. So if I am right that Twin Earth presents no real challenge, then there is no need to define a notion of narrow content in response. So neither Fodor nor anyone else needs narrow content. Narrow content is a specious notion constructed only out of a desperate need to solve the Twin Earth problem – but this problem can be dissolved before we need to postulate narrow content.

Of course, if narrow states are just, by definition, states that any pair of ‘Twins’ share, then I accept that intentional states are narrow states. But this does not mean that these states do not have content in the ordinary, truth-conditional sense. It is just that any difference in the contents of these states must be mediated by a difference in their intrinsic properties. To put it in terms reminiscent of Fodor’s formality condition: there is no difference in content without some intervening difference in intrinsic properties.

So if intentional states are to be the causes of behaviour, they have to cause behaviour indirectly, via some intrinsic property of the thinker. Fodor agrees (e.g., 1981, pp. 201–2), and he also agrees that these intrinsic properties – ‘mental representations’ – must be *correlated* with the contents of these states, to give them the right causes and effects. For instance, the intrinsic properties of my belief that water is wet must be correlated (somehow!) with the content *water is wet* in order for it to cause me to ask for water and not whisky at the bar. And the intrinsic properties of my desire for water must be correlated (somehow!) with the content *I drink some water* in order for my action – asking for water – to result in the satisfaction (or not) of *this* desire.

To say how thinkers' intrinsic properties are correlated with the contents of their thoughts would be to solve the problem of intentionality. Unfortunately, I do not have a solution to this problem. But if what I say here is right, then the Twin Earth argument has to be dissolved before the problem of intentionality is solved. This is because the Twin Earth argument entails either that intentional states are not the causes of behaviour; or that the contents of intentional states can affect what they cause without any intrinsic causal mechanism. If the first of these is true, then there really is no problem of intentionality; and if the second is true, then the problem of intentionality is solved by something like an appeal to magic. I think neither of these are satisfactory solutions to the problem: dissolving Twin Earth is the only option.

So Putnam and Burge have not shown intentional states to be broad. Burge's twins differ only in the truth values of their beliefs about the right word to use in their communities; Putnam's twins may differ in this way too (cf. 'aluminium') or they only differ according to an essentialist re-description of the contents of their thoughts (cf. 'water'), a description that tells us nothing about psychological reality, since the alleged difference in no way affects their relevant intrinsic properties. The first difference depends on a difference in communal linguistic practice, but does not show much about intentional states; the second alleged difference is the consequence of an undefended and implausible essentialism. So the path is cleared for the natural, obvious conclusion: what the twins have in common is a state with ordinary, truth-conditional content: nothing more, nothing less. All the differences are in the world.²²

University College London

REFERENCES

- Almog, J., Perry, J. and Wettstein, H. (eds) (1989) *Themes From Kaplan* (Oxford: OUP).
 Bach, K. (1982) 'De Re Belief and Methodological Solipsism', in Woodfield (1982).
 — (1988) 'Burge's New Thought-Experiment: Back to the Drawing Room', *Journal of Philosophy*, 85 pp. 88–97.
 Blackburn, S. (1984) *Spreading the Word* (Oxford: Clarendon Press).
 Block, N. (ed.) (1981) *Readings in the Philosophy of Psychology*, 2 vols (London: Methuen).
 — (1987) 'Functional Role and Truth Conditions', *Proceedings of the Aristotelian Society Supplementary Volume*, 61 pp. 157–81.
 Burge, T. (1978) 'Belief and Synonymy', *Journal of Philosophy*, 75 pp. 119–38.
 — (1979) 'Individualism and the Mental', *Midwest Studies in Philosophy* vol. IV, pp. 73–121.

²² I am grateful to the participants at a meeting of the Wolfson Philosophy Society in Oxford, and to Tom Baldwin, Jeremy Butterfield, Naomi Eilan, Melinda Hogan, Greg McCulloch, Hugh Mellor, Mark Sainsbury and especially Gabriel Segal for discussions that have helped me with this paper.

- (1982) 'Other Bodies', in Woodfield (1982).
- (1986a) 'Individualism and Psychology', *The Philosophical Review*, 95 pp. 3–45.
- (1986b) 'Cartesian Error and the Objectivity of Perception', in McDowell and Pettit (1986).
- (1986c) 'Intellectual Norms and the Foundations of Mind', *Journal of Philosophy*, 83 pp. 697–720.
- Butterfield, J. (1986a) 'Content and Context', in Butterfield (1986b).
- (ed.) (1986b) *Language, Mind and Logic* (Cambridge: Cambridge University Press).
- Carruthers, P. (1987) 'Russellian Thoughts', *Mind*, 96 pp. 18–35.
- Crane, T. (1990) 'The Language of Thought: No Syntax Without Semantics', *Mind and Language*, 5.
- and Mellor, D.H. (1990) 'There is No Question of Physicalism', *Mind*, 99 pp. 185–206.
- Davidson, D. (1984) *Inquiries into Truth and Interpretation* (Oxford: Clarendon).
- (1987) 'On Knowing One's Own Mind', *Proceedings and Addresses of the American Philosophical Association*
- (1988) 'Reply to Burge', *Journal of Philosophy*, 85 pp. 664–5.
- Dennett, D. (1982) 'Beyond Belief', in Woodfield (1982).
- Dupré, J. (1981) 'Natural Kinds and Biological Taxa', *The Philosophical Review*, 90 pp. 66–90.
- Evans, G. (1982) *The Varieties of Reference* (Oxford: Clarendon Press).
- Field, H. (1972) 'Tarski's Theory of Truth', *Journal of Philosophy*, 69 pp. 347–75.
- (1978) 'Mental Representation', reprinted in Block (1981).
- Fodor, J. (1975) *The Language of Thought*, (New York: Crowell).
- (1980) 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology', repr. in Fodor (1981).
- (1981) *Representations* (Hassocks: Harvester Press).
- (1987) *Psychosemantics* (Cambridge, Mass.: MIT Press).
- Geach, P.T. (1969) *God and the Soul* (London: Macmillan).
- Jackson, F. and Pettit, P. (1988) 'Functionalism and Broad Content', *Mind*, 97 pp. 381–400.
- Kaplan, D. (1977) 'Demonstratives', in Almog, Perry and Wettstein (1989).
- Kripke, S. (1980) *Naming and Necessity* (Oxford: Blackwell).
- Lewis, D. (1983) 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, 61 pp. 343–77.
- McCulloch, G. (1986) 'Scientism, Mind and Meaning' in McDowell and Pettit (1986).
- (1989) *The Game of the Name* (Oxford: Clarendon Press).
- McDowell, J. (1984) 'De Re Senses', *The Philosophical Quarterly*, 36 pp. 283–94.
- (1986) 'Singular Thought and the Extent of Inner Space', in McDowell and Pettit (1986).
- and Pettit, P. (eds) (1986) *Subject, Thought and Context* Oxford; Clarendon Press.
- McGinn, C. (1977) 'Charity, Interpretation and Belief', *Journal of Philosophy*, 74 pp. 521–35.
- (1982) 'The Structure of Content', in Woodfield (1982).
- (1989) *Mental Content* (Oxford: Blackwell).
- Mellor, D.H. (1977) 'Natural Kinds', *British Journal for the Philosophy of Science*, 28 pp. 299–312.
- (1981) *Real Time* (Cambridge: Cambridge University Press).
- (1989) 'I and Now', *Proceedings of the Aristotelian Society*, 89 pp. 79–94.
- Noonan, H. (1984) 'Fregean Thoughts', *The Philosophical Quarterly*, 36 pp. 205–24.
- (1986) 'Russellian Thoughts and Methodological Solipsism', in Butterfield (1986b).
- Perry, J. (1979) 'The Problem of the Essential Indexical', *Nous*, 13 pp. 3–21.
- Putnam, H. (1975) *Mind, Language and Reality* (Cambridge: Cambridge University Press).
- (1988) *Representation and Reality* (Cambridge, Mass.: MIT Press).
- Quine, W. (1960) *Word and Object* (Cambridge, Mass.: MIT Press).
- Searle, J. (1983) *Intentionality* (Cambridge: Cambridge University Press).
- Segal, G. (1989a) 'The Return of the Individual', *Mind*, 98 pp. 39–57.
- (1989b) 'Seeing What is Not There,' *The Philosophical Review*, 97 pp. 189–214.
- Shoemaker, S. (1984) *Identity, Cause and Mind* (Cambridge: Cambridge University Press).
- Woodfield, A. (ed.) (1982) *Thought and Object* (Oxford: Clarendon Press).