

PIERRE JACOB

EXTERNALISM REVISITED:
IS THERE SUCH A THING AS NARROW CONTENT?

(Received 31 May, 1989)

There are presently, or so it seems to me, three main views about the individuation of the content of intentional mental states. First, there is the *no* content view: the view that the content (of an intentional mental state) is an ordinary folk pre-scientific concept which can play no role in a respectable mature scientific psychology. Second there is a *dualist* (or bifurcationist) view of content which posits a distinction between narrow and broad content. Although it acknowledges the (obvious) existence of broad content and because it assumes the existence and legitimacy of narrow content, dualism is consistent with an individualist view of content. Third, there is anti-individualism, which is a *monist* view about the individuation of the content of intentional mental states such as thoughts, beliefs and other propositional attitudes, i.e. states which we ordinarily or commonsensically describe or ascribe to a person by means of a complex sentence with a psychological verb (describing the person's attitude) prefixed by a singular term referring to the person and followed by an embedded clause (or "that"-clause) expressing the content of the person's attitude. Unlike dualism, anti-individualism assumes that there is no room for a narrow individuation of the content of an individual's intentional mental states since it cannot ever be individuated independently from the individual's (physical or social) environment.

According to anti-individualism, a person's brain, his or her brain states and (probably) his or her "subdoxastic" cognitive states are located within the person's cranial box.¹ Many if not all of the person's thoughts and intentional mental states however would be spread unto the person's physical and social environment on which they depend for their individuation. According to anti-individualism, a person's physical and social environment is inherently part of the content of his or her thoughts and other intentional mental states. So if anti-individualism is

right, then what Fodor (1980, p. 229) refers to as “the long tradition, including both Rationalists and Empiricists, which takes it as axiomatic that one’s experience (and a fortiori, one’s beliefs) might have been just as they are even if the world had been quite different from the way it is” is just wrong.

There are at least two versions of, or two main roads leading to, anti-individualism depending on whether one focuses upon the contribution of particular (“physical”) objects located in a person’s (“physical”) environment to the content of his or her so-called “singular thoughts” or upon the contribution of a person’s (social and) linguistic community to the content of his or her more general thoughts. The first road is roughly to adapt the “theory of direct reference” initially proposed by Donnellan, Kripke, and Kaplan for the analysis of the (“Russellian”) contents of utterances (of sentences containing singular terms) to the contents of thoughts and intentional mental states (thereby expressed). Such is the strategy of Evans (1982) and McDowell (1977) when they argue that the content of what they call a “singular thought” depends for its very existence (or “availability”) on the existence of the object to which the thought applies: on their view, unless the object of a purported singular thought exists or unless the singular term contained in the sentence expressing the thought and purportedly used to refer to the object has a reference, the purported thought fails to qualify as a genuine singular thought — if not as a thought at all. This is not the road chosen by Tyler Burge in his argument for social anti-individualism and I will not examine it here. Burge’s anti-individualism can be seen as a possible conclusion derived from an argument constructed out of a series of thought-experiments the first of which perhaps goes back to Putnam’s famous (1974) Twin Earth thought-experiment — where the argument relies essentially upon a semantic analysis of the truth-conditions of “that”-clauses used to ascribe beliefs and other intentional mental states.

In the present paper, I want to propose an alternative to Burge’s anti-individualist interpretation of his important thought-experiments, an alternative which is closely akin to the individualist response provided by Brian Loar (1987a, b) and which I take to be consistent with the *dualist* view of content which has been espoused by a number of

philosophers. In the first two sections of the paper, I will briefly examine a number of possible reactions to Putnam's initial Twin Earth thought-experiment. In particular, I want to examine the possible defense of content-dualism drawn from the parallel between the contents of mental states and the meanings of words in natural languages. In the third section, I will lay out as explicitly as I can Burge's argument for anti-individualism and for his rejection of content-dualism based on thought-experiments of his own. In the fourth section, I will cast doubt on Burge's premisses. In the fifth and last section, I will accept Burge's conclusion and I will try to argue that his monist anti-individualist view of content still requires a further premiss which I examine critically. I will then provide an individualist alternative to the premiss in question and sketch my own dualist interpretation of Burge's thought-experiments which will be based on the innocuous assumption that a speaker's utterance is meant to be an interpretation of his or her thought.

I. HOW TO DERIVE THE THREE VIEWS ABOUT CONTENT FROM TWIN EARTH?

As is well known, Putnam (1974) imagined that on Twin Earth the same word "water", which is used on Earth to refer to the liquid with molecular structure H_2O , is used to refer to a different liquid with many similar phenomenological properties but with a different molecular structure. So we are to imagine that for each and every of my utterances of a sentence containing the word "water" on Earth, my twin (or my doppelgänger) utters a different token of the same sentence-type containing the word "water". Two such simultaneous utterances of the same word-types will have different truth-conditions for the word "water" on Earth and the word "water" on Twin Earth do not have the same conditions of satisfaction. We may suppose that the fact that, unlike my twin's, my brain might contain H_2O molecules (or that it is the brain of a person whose body contains H_2O molecules) is irrelevant to the identity of my twin's and my brain states.² If so, we may assume that my twin and I are in relevantly identical brain states, identical internal physical and functional states, that we have the same behavioral dispositions, that the histories of our respective interactions with proximal

stimuli are the same — although some of the distal causes of our mental lives and behaviors (e.g. H₂O molecules as opposed to some other chemical compound) are different.

To many philosophers, principle (a) has seemed like a natural constraint on any plausible concept that deserves to be called content and principle (b) has seemed like a reasonable condition on any “scientific” approach to the content of mental states:

- (a) content is or should determine truth-conditions;³
- (b) individuals in the same “physical” condition (in the same brain state) should be said to be in a psychological (or mental) state with the same content;

or equivalently

- (b’) the content of the psychological (or mental) state of a person should supervene on his or her “physical” (brain) state.

From Putnam’s Twin Earth thought-experiment together with premisses (a) and (b), the three competing views about the individuation of the contents of mental states which I alluded to at the beginning can be derived: the no content view, the dualist view of content and the monist anti-individualist view of content.

First it may be and has been argued (most notably by Stich 1978a, 1983, 1988, who calls the argument “the Autonomy argument”) that the thought-experiment jeopardizes the commonsensical view that thoughts, beliefs and other propositional attitudes play a role in the explanation of human action and behavior in virtue of their contents. Let us ask: does one member of a pair of Putnamian doppelgängers believe the same thing as the other member of the pair when each utters the sentence “The water in my bath-tub is too hot”? In view of the fact that their two simultaneous utterances have different truth-conditions (by maxim (a) narrowly or severely construed as asserting the identity of content with truth-conditions), it is tempting to give a straightforward negative answer. But by assumption the doppelgängers are in relevantly identical brain states. So it seems to follow that the property of *believing that the water in one’s bath-tub is too hot* does not supervene on a person’s (the speaker’s) brain state or internal physical

constitution. If now we combine this proposition with the plausible view (maxim (b)) that the psychological explanation of a person's action and behavior does or should appeal only to states or properties of a person which do supervene on the person's internal physical and functional condition, we reach the conclusion that a property such as *believing that the water in one's bath-tub is too hot* cannot be invoked in a psychological explanation of why each member of the pair of Putnamian doppelgängers did select the same word-types to express their respective thought and to predict the fact they are both likely to turn on (the faucet for) cold water. Stich has exploited this argument to cast doubt on the explanatory utility and scientific respectability of content.

Before I examine Burge's own reaction to Putnam's thought-experiment, I want to pause briefly to emphasize that Putnam himself suggested two views one might take: a monist anti-individualist view pursued by Burge and a dualist view.

One may argue that the difference in the extension of an utterance of "water" on Earth and on Twin Earth counts as a difference in meaning. Let us call this the "thick" sense of the meaning of "water". On the thick sense, "meaning is not in the head" and none of an Earthian "water"-utterances on Twin Earth could express a true proposition. This is the monist anti-individualist line taken by Burge — I shall return to it momentarily.

What motivates the content-dualist reaction is that, although the word "water" does not have the same extension on Earth and on Twin Earth, it is natural to describe, as Putnam (1974, p. 224) does, the pair of doppelgängers as "exact duplicates in appearance, feelings, thoughts, interior monologue etc.". If so, the thought-experiment suggests that there must be a non truth-conditional notion of content. The dualist reaction has been accepted by a number of philosophers who might otherwise disagree on much else (Bach 1982, Block 1986, Dennett 1982, McGinn 1982, Field 1977, Fodor 1987, Harman 1982, Kaplan 1977, Loar 1987a, b, Perry 1977) and it has been expressed in terms of a distinction between *broad* (or *wide*) content and *narrow* content. The former has been linked to the truth-conditions of a thought, the latter to the role of the thought in deliberation and the planning of action (what is also called the thought's "functional" or "conceptual role"). So the content-dualist would say that by their "water"-utterances

each of a pair of Putnamian doppelgängers expresses a thought with the same narrow content and different broad contents.

Rather than emphasizing the fact that the “thick” meaning of “water” is not the same on Earth and on Twin Earth, one may also argue, as Putnam did, that “water” is an *indexical* word or has a hidden indexical component. On the one hand an indexical word like “I”, “now” or “here” has a constant linguistic meaning tacitly known to any competent speaker or hearer of the language — what Kaplan calls the “character” of the word. In this “thin” sense of the word, linguistic “meaning” is just what is tacitly known to any competent user of a word in virtue of which he or she is able to use the word appropriately. It is not what allows any competent speaker to pick up unerroneously any correct instance of the extension of the word. In the case of a natural kind word, linguistic meaning must be supplemented by specialized knowledge of the world. In the case of demonstratives, it may have to be supplemented by assumptions about the speaker’s intentions (as in the case of a token of “he”). On the other hand “I”, “now” or “here” can be used to refer to different persons, different instants of time or different places on different occasions of use. So indexical words make different contributions to the propositional content they are used to express on different occasions. To claim that “water” is indexical amounts to assuming that the “thin” linguistic meaning of “water” (what Putnam calls the “stereotype” associated with the word) is constant and the propositional contribution made by “water” changes as the word is uttered on Earth or on Twin Earth.

II. WHY CAN'T NARROW CONTENT BE LINGUISTIC MEANING?

So the content-dualist may rightly be asked: What is the narrow content of a “water”-utterance such that, in spite of the fact that “water” refers to H₂O on Earth, but not on Twin Earth, two members of a pair of Putnamian doppelgängers can be said to express the same narrow thought-constituent by means of their respective utterance of the word “water” on Earth and on Twin Earth? It is tempting for a content-dualist to respond that the narrow content in question is just the linguistic meaning of the indexical word “water”. In other words, it is tempting for the content-dualist to identify, as Kaplan (1977) and Perry

(1977, 1979) have done, the broad content of a thought with the truth-conditions of the utterance (its “thick” meaning) used to express it and the narrow content of the thought with the linguistic meaning of the (indexical) sentence uttered, as represented in Diagram 1:

UTTERANCES	THOUGHTS
linguistic meaning	narrow content
_____	_____
truth-conditions	broad content

DIAGRAM 1

There are at least two objections to this dualist interpretation of Putnam’s thought-experiment. One may, as Burge (1982) does, object to the claim that “water” is an indexical word on the ground that with a genuine indexical word like “here”, I may on different occasions refer to two different places and say something *true* on both occasions. If “water” were indexical, it would have a constant linguistic meaning and would change reference on different occasions. So were I to visit Twin Earth, by my use of the word “water”, I could say something true. But according to Burge (1982, pp. 104–105), who accepts Putnam’s other suggestion that extension is part of meaning and that “meanings are not in the head”, I would thereby express a *falsehood* since there is no water on Twin Earth. Adjudication between the two views is no easy task. I will here merely record the fact that there are two possible rival views one can take to describe the situation: the view that by his “water”-utterance on Twin Earth, an Earthian would automatically say something false since there is no water there; and the view that although “water”-utterances do not express the same propositions on Earth and on Twin Earth, an Earthian utterance of a sentence containing the word “water” is not automatically false; it may be true. The former view entails that “water” on Earth and “water” on Twin Earth are homonyms. The latter is compatible with the claim that “water” is indexical. Actually, the former view cannot be correct for suppose that, were an

Earthian to utter "This is water" on Twin Earth, his utterance would be false; however had he instead uttered "This is *not* water", his utterance would automatically be true insofar as it is the negation of a falsehood.

The second objection which seems more serious to me (and which has been expressed among others by Loewer & Lepore 1986, Recanati 1989 and Wettstein 1986) is that the narrow content of a thought cannot be identified with the linguistic meaning of a sentence containing an indexical or demonstrative expression. I will give three reasons for rejecting such an identification, the first two being closely related.

First, the linguistic meaning of the first person indexical pronoun "I" is roughly the rule: 'By any token of "I" the speaker refers to himself or herself'. So by virtue of its linguistic meaning (tacitly known to speaker and hearer), "I" presents its referent as the speaker to both speaker and hearer. But it is plausible enough that the speaker, who thinks of himself or herself in a first person way, does not think of himself or herself as the speaker or under the description *the person who uttered this token of "I"*; nor does the speaker think of himself or herself in the same way as the hearer thinks of the speaker since the hearer, unlike the speaker, might think of the speaker as the speaker. So the narrow thought-constituent expressible by "I" cannot be identical to the linguistic meaning of "I".

Second, the linguistic meaning of the second person pronoun "you" is roughly for both speaker and hearer: 'By any token of "you" the speaker refers to the addressee or the person to whom the utterance is addressed'. The speaker's narrow thought-constituent expressed by "you" might be identical to the linguistic meaning of "you", since the speaker might think of his or her hearer as his or her addressee. But the addressee will, upon understanding the speaker's utterance, presumably think of himself or herself in a first person way, not at the addressee of the speaker's utterance.

Finally, to take an example of Evans's (1982, p. 84), consider Alfred, a rational speaker not prone to expressing contradictions who utters "This ship is a Japanese steamer but this ship is not a Japanese steamer", where he takes himself to be watching two different ships seen through two different windows, whereas unbeknownst to him, he has referred twice to the same ship. The two occurrences of "this ship" cannot have the same "cognitive significance" or cannot express the same narrow

thought-constituent since to assume otherwise would be inconsistent with our assumption that Alfred is (minimally) rational (not prone to giving voice to contradictory thoughts or beliefs): if Alfred had realized that his two tokens of “this ship” referred to the same ship, he would not have uttered the words he uttered. Perhaps — assuming the ship to be descriptively identified — the narrow content of the thought expressed by the first occurrence of “this ship” in Alfred’s utterance might be *the ship visible by the window to my left* while the narrow content of the thought expressed by the second occurrence of the same words might be *the ship visible to my right*.⁴ However the linguistic meaning of the demonstrative expression “this ship”, unlike the above narrow content of the thought expressed by each token, does not change from one occurrence to the next. Nor is the meaning of “this ship” identical to either sense. So the narrow content of a thought cannot be identical to the linguistic meaning of indexical words used to express it.

The following conclusion seems therefore inescapable: the narrow content of a thought *cannot* be identified with the linguistic meaning of indexical words used to express it. But it is incumbent upon the content-dualist to say what narrow content might be. Does not the above conclusion refute content-dualism? Does not it justify the anti-individualist claim that narrow content ought to be called “content by courtesy” (cf. Baker 1987)? Before I try to justify a negative response to this question, I want to devote the next section to scrutinizing Burge’s argument for his monist anti-individualist view of content. My response to Burge will provide an appropriate background for my justification of the claim that content-dualism is not refuted by acknowledging that the narrow content of a thought cannot be identical to linguistic meaning.

III. BURGE’S ARGUMENT FOR ANTI-INDIVIDUALISM

First, as we saw in the previous section, Burge denies that “water” is indexical. For him “water” on Earth and “water” on Twin Earth are two homonyms with two different meanings, two different word-types. Perhaps when reporting in our Earthian language what my twin said or thought on Twin Earth by his utterance of “water” I ought to use “twater”, a word invented for the purpose of reporting in my language

utterances of, and thoughts expressed by, “water” on Twin Earth. But as will become clear later, I do not agree with this suggestion. Burge claims that no sense can be made of the dualist assertion that two members of a Putnamian pair of doppelgängers can be said to express a thought with the same narrow content by means of their “water”-utterances on Earth and on Twin Earth. As Burge (1982, p. 109) says:

There is no water on Twin Earth, so [a Twin Earthian] has never had any contact with water. Nor has he had contact with anyone else who has had contact with water. Further, no one on Twin Earth so much as used a word which means *water*. It is not just that water does not fall in the extension of any of the Twin Earthian’s terms. The point is that none of their terms even translates into our (non-indexical) word ‘water’. No English_{TE}-to-English dictionary would give ‘water’ as the entry for the Twin Earthians’ word. It would thus be a mystery how a Twin Earthian could share any of [an Earthian’s] attitudes that involve the notion of water. They have not had the normal means of acquiring the concept. The correct view is that they have acquired, by entirely normal means, a concept expressed in their language that bears some striking, superficial similarities to ours. But it is different.

One of Burge’s two reasons for rejecting the dualist view that “water” can be used on Earth and on Twin Earth to express a thought-constituent with the same narrow content is his already noted critique of the claim that “water” is an indexical word. Burge’s other reason is that causal interaction with two different chemical compounds cannot result in identical thought-constituents (or mental representations). To which the content-dualist would of course respond that representations with different causes might have different broad contents and the same narrow content.⁵

In the present section, I want to lay out Burge’s (1979) important argument for anti-individualism. To do this, I will first review briefly two of Burge’s own fascinating thought-experiments. Then I want to provide some motivation for weakening the force of the argument. In the next section, I will accept the argument as valid and try to show how its conclusion can be reconciled with individualism. This done, I will explain why I believe that acknowledging that narrow content is not identical to linguistic meaning does not refute content-dualism.

Burge’s arthritis example. Let us consider Alfred who suffers from, and who thinks that he suffers from, arthritis. He has several correct beliefs about his medical condition: he believes that his arthritis in his fingers is more painful than his arthritis in his ankles, that his father too had

arthritis, that it is better to have arthritis than to have a heart-attack. But he also has some incorrect beliefs: one morning, he wakes up with a pain in his thigh which reminds him of his other rheumatoid pains and he thinks erroneously that his arthritis has spread into his thigh. So he consults his doctor who corrects him and informs him that he cannot have arthritis in his thigh since arthritis is a disease of the joints, not muscles. So he is relieved and he gives up his erroneous belief.

Let us pause to emphasize the fact that on Burge's view (which I assent to), in spite of his misunderstanding about what arthritis really is (or in spite of his inaccurate conception of arthritis), Alfred expresses beliefs about arthritis by his use of the word "arthritis" and beliefs can be ascribed to him by an ascriber using the word "arthritis".

Let us now imagine Alfred with the same internal constitution and the same biography as above living in a hypothetical community members of which speak a Burgian dialect of English which differs from actual English merely in that "arthritis" does not have exactly the same meaning as in our community: in the counterfactual community, it refers not to a disease of the joints but to various rheumatoid ailments located in joints, tissues, bones and muscles. Everything about Alfred (and his fellows) — his chemical, neurological, physiological, functional and behavioral history — that can be described in non-intentional terminology is the same in both communities. When in the counterfactual community, Alfred wakes up one morning thinking that he has arthritis in his thigh (or should *I* say "thinking that he has tharthritis in his thigh"? — a thought *he* would express using the word "arthritis"), he decides to pay a visit to his doctor and he expresses his fear that his arthritis/tharthritis has spread into his thigh using the exact same word form which he uses in the actual community. However in the counterfactual community, the doctor corroborates his assumption that he has arthritis/tharthritis in his thigh. When Alfred utters the sentence-type "I now have arthritis in my thigh" does he or does he not express the same belief in both communities? Burge answers: "No".⁶

Burge's sofa example. We imagine that Alfred has several correct and a few incorrect beliefs about sofas: e.g. he may correctly believe that sofas are pieces of furniture on which people sit, that they are couches; and he may also incorrectly believe that broad single-seat armchairs are

sofas (which they are not). By the maxim alluded to earlier, although Alfred has not entirely mastered the meaning of “sofa” (or has an inaccurate conception of what sofas really are), we would still, by Burge’s light, credit him with the concept of a sofa; he would express beliefs about sofas by uttering the word “sofa” and we might ascribe to him beliefs about sofas using the word “sofa”.

If we now imagine Alfred with no internal change in a counterfactual Burgian community whose members speak a language in which “sofa” refers to broad single-seat armchairs as well as to couches, we should not, by Burge’s light, credit Alfred with the concept of a sofa; and we should neither assume that he expresses beliefs about sofas when he utters the word “sofa” nor should we ascribe beliefs to him about sofas using our word “sofa”. Perhaps, as argued in Burge (1986b), should we invent the word “safo” to so report Alfred’s thought, relative to the hypothetical community.

If we accept Burge’s verdict on the facts (both actual and hypothetical), then what I take to be Burge’s (1979) main argument for anti-individualism can now be set up in explicit form:

- (1) “Alfred believes that he has arthritis in his thigh” can be truly used to ascribe a belief to Alfred relative to the actual community;
- (2) “Alfred believes that he has arthritis in his thigh” cannot be truly used to ascribe a belief to Alfred relative to the counterfactual community;
- (3) Alfred is in an identical internal “physical” state in the actual and in the counterfactual community;
 Subconclusion: belief-ascriptions of the form “Alfred believes that . . .” do not merely describe internal states of the believer or are not made true merely by internal states of the believer;
 CONCLUSION C: the truth-conditions of belief-ascriptions are partly constituted by facts pertaining to the environment of the person who is being ascribed a belief.

Although the argument is, I believe, a sound one, I note that a theory that does not do justice to the intuitive *similarities* between Alfred’s mental state in either community when he wakes up one morning

thinking and fearing that his arthritis has spread into his thigh seems quite implausible to me. A better theory would account for both our inclination to say in some circumstances that Alfred has different beliefs in the two communities *and* for our inclination to say in other circumstances that he has the same belief in the two communities. As Burge (1979, pp. 74–77 and 1982 *passim*) insists, to evaluate the argument, we must restrict ourselves to *de dicto* belief-ascriptions, i.e. to attempts on the part of the ascriber to capture the content of the believer's point of view by his or her use of relevant terms in oblique position embedded in "that"-clauses; we do not consider *de re* belief-ascriptions in which an ascriber can actually use a singular term not to try and capture the content of the believer's point of view but to refer to what the singular term refers to. However conclusion *C* does not assert that the (social) environment is a constitutive part of the contents of a person's thoughts or beliefs. The further conclusion that the contents of a person's thoughts cannot be individuated independently from the person's environment has not yet been reached. In order to infer the latter conclusion, one has to identify the truth-conditions of a belief-ascription with the content of the thought or the belief ascribed. This further step — which I take to be the central issue — I will question in the last section. For now, I want to consider reasons to mitigate acceptance of the first two premisses of the above argument.

There are two differences between Putnam's and Burge's thought-experiments which might be thought to indicate the superiority of Burge's thought-experiments over Putnam's as a fuel for anti-individualism. First Burge does not rely on the supposition of there being two doppelgänger since we imagine the very same individual in two communities. Second, whereas Putnam's example seemed restricted to natural kind words, Burge's thought-experiment seems applicable to *any* word in the language.⁷ So Putnam's thought-experiment can only be used by the anti-individualist to argue that the "physical" environment is part of the content of an individual's thought when the thought is about a fragment of his or her "physical" environment — a conclusion perhaps reminiscent of Evans and McDowell's view about what makes a singular thought "available". Burge's thought-experiment can be used to show that an individual's *social* environment is part of the content of any of the individual's thought.

IV. CASTING DOUBT UPON BURGE'S PREMISES

Now let me try first to weaken the strength of the anti-individualist argument by turning the apparent strength of Burge's examples against it. What I want to do at present is question the asymmetry involved in the joint acceptance of premisses (1) and (2). I want to suggest a kind of parity: if a belief can be truly ascribed to Alfred in the actual community using certain words in spite of his own inaccurate understanding of the meaning of those very words, then there are grounds for ascribing to him, relative to the counterfactual community, a belief using the same words in spite of differences in the communal meaning of those words. Conversely, our grounds for doubting that we can ascribe a belief to Alfred, relative to the counterfactual community, ought to be grounds for refusing to ascribe to him a belief, relative to the actual community, using the same words when, relative to our actual community, Alfred's conception of arthritis betrays some misunderstanding. Besides the pragmatics of belief-ascriptions, which does I think support parity, I want to point out a deep difference between Burge's and Putnam's thought-experiments and then I want to scrutinize Burge's insistence on the role of an individual's disposition to defer to the opinion of expert members of his community for the individuation of the contents of his thoughts. Notice that the pragmatic factors I am about to consider are properties of *de dicto*, not *de re* belief-ascriptions.

I hasten to add that I accept the principle that Alfred's inaccurate conception of arthritis should not prevent us from ascribing to him beliefs using the word "arthritis" relative to our actual community. How else could we ascribe a belief to him? How else could we describe his mistake? Therefore by my parity maxim, I accept it that the same word form ought to be usable to ascribe beliefs to him relative to the counterfactual community. Unless perhaps Alfred turns out to have an expert understanding of "arthritis"; only he thinks that the present conception of arthritis as developed in our present community is inadequate and he assumes that he has good reasons to believe that arthritis is a disease that can occur concurrently in joints, tissues and muscles.

IV.1. *A Pragmatic Argument*

Suppose I am talking to Paul. We are both acquainted with Alfred and

with the (Burgian) peculiarities of Alfred's linguistic community in which "arthritis" is used to refer to rheumatoid ailments in muscles, tissues and joints. I want to convey to Paul Alfred's fear that his arthritis has spread into his thigh — a fear that would be relieved in our community but will not be in Alfred's community. So I use the words "Alfred thinks he has got arthritis in his thigh", although I know and I know that Paul knows (and I know that he knows that I know . . .) that "arthritis" does not in Alfred's community refer exactly to what it refers to in our community. Notice I am not producing a *de re* belief-ascription. I choose the word "arthritis" because it is the best word in Paul's and my language to try and capture Alfred's fear (his own personal or subjective perspective on his condition at a certain stage in his life) in spite of (or given) the difference between his or our language. It is up to the context in which my belief-ascription takes place to fill in the missing words: my use of the word "arthritis" in the belief-ascription might be supposed to be short for: "what in Alfred's community they take arthritis to be" or something of that kind. The fact that my utterance *abbreviates* my thought, I take to be crucial and will try to account for in the next section. I could in this case add further words to my report of Alfred's state of mind depending on the amount of information about Alfred's community mutually known to me and my audience. But I may rely on context.

This is a widespread linguistic phenomenon — noted by Donnellan (1966) — which can be illustrated by facts simpler than belief-ascriptions. Suppose that Paul and I are at a party and we meet Edith in the company of John whom Paul mistakenly takes to be Edith's husband whereas John is actually Edith's lover. If Paul's belief that John is Edith's husband is mutually known to Edith and me, then upon noticing the fact that John is walking towards us, I may use the sentence "Your husband is coming" to communicate to Edith the proposition that John is coming. I might have thereby said something true of John by referring to him under the description "your husband". It is then plausible to maintain that by using "your husband" I merely "introduce" John (Edith's lover, not Michael, Edith's husband) into the proposition expressed.⁸ Arguably, the definite description "your husband" would be short for the longer description "your husband according to Paul" or "the man Paul takes to be your husband". Again a case of a thought summarized by an utterance. Given its conventional linguistic meaning,

the word “husband” can be part of a definite description used to refer in most contexts to Edith’s actual husband, Michael, and in some contexts to John, Edith’s lover. I can use the word “husband” to refer (directly or indirectly, by means of a reference to Paul’s system of beliefs) to Edith’s lover. Although such uses of words are perfectly ordinary, we might rule them out as improper, idealize away from them and claim that they require some special treatment, on the ground that the concept expressed by the word (here, the concept of being Edith’s husband) has been improperly invoked. But then it would be odd to claim that we *can* use the word “arthritis” which refers to a disease of the joints to ascribe *de dicto* a belief to Alfred in our actual community when it is mutually known to the belief-ascriber and his audience that Alfred’s conception of arthritis is inaccurate.

My previous remarks have an obvious bearing on the truth-conditions of “arthritis”-sentences uttered in the counterfactual community as we assign them in our actual community. As already noted, accepting the truth of premiss (2) amounts to assuming that “tharthritis”, not “arthritis”, ought to be used to report in my language Alfred’s thought, relative to the counterfactual community. So that on Burge’s view, the truth-condition of Alfred’s utterance, relative to the counterfactual community, of “I now have arthritis in my thigh” is:

“I now have arthritis in my thigh” is true iff Alfred has tharthritis at the time of utterance.

Where “arthritis”, a word of the counterfactual community, is being mentioned and “tharthritis” is being used as a word of my language operating as the metalanguage. Presumably, on Burge’s view, “tharthritis” has an empty extension since according to experts of my (metalinguistic) community, there is no disease such that rheumatoid ailments can jointly occur in joints *and* muscles *and* tissues. So Burge’s view entails that none of Alfred’s thoughts, relative to the counterfactual community, expressed by an utterance of “arthritis”, is true — since nothing, by the lights of the actual community, satisfies his use of “arthritis”: there is no disease that occurs concurrently in joints *and* muscles. We, members of the actual community, ought to take the word “arthritis” as used in the counterfactual community as a term as devoid of reference as we take the word “phlogiston” to be. I think this simply is the homonymy view

alluded to earlier with respect to “water”. But it is not inevitable. As I may take “arthritis” used in the counterfactual community to have a *partial* extension, I may take *some* of Alfred’s thoughts, relative to the counterfactual community, expressed by means of his utterance of “arthritis” (such as “my father had arthritis”) to be *true* and assign to his utterance of “I now have arthritis in my thigh” the following truth-condition:

“I now have arthritis in my thigh” is true iff Alfred has arthritis at the time of utterance.

Just as in the preceding statement of truth-condition, both sides of the biconditional being false, the biconditional is true. I would argue that were experts of the two communities to get together, experts from the counterfactual community would give up their conception of arthritis and accept ours.

IV.2. *Why Do “Arthritis”-Utterances Have the Same Truth-Conditions in Both Communities?*

Burge’s asymmetric treatment of ascriptions according to whether Alfred belongs to one or the other community suggests that Alfred can no more be credited with the same concept of arthritis in both communities than members of a pair of Putnamian doppelgängers could be credited with the same concept of water.

Now I want to resist this suggestion by noticing that “arthritis” is the name of a disease. So we expect bio-medical experts in either community to uncover truths about the causes and etiology of arthritis — truths which depend on physiological non-intentionally describable facts of the matter. On Burge’s own hypotheses, Alfred’s physiological condition does not change from one community to the next. Only experts’ opinions about the disease vary from one community to the next. Presumably experts’ opinions about arthritis do not “create” the physiological (constant) facts — they are merely representations made true or false (correct or incorrect) by (constant) physiological facts. So it makes sense to assume that one community has the facts right and the other wrong and that all things being equal, and given enough time, experts from both communities will or would converge on their concep-

tion of arthritis. Unlike Putnam's example in which "water"-utterances on Earth and on Twin Earth have different truth-conditions (due to differences in the chemical environment), "arthritis"-utterances in both communities should be ascribed the same community-independent truth-conditions (due to constant physiological conditions in both communities). So the natural thing to say is that although the two communities have developed two *conceptions* of arthritis, Alfred may be credited with the same *concept* in both communities: his mental carving up of the phenomenon is the same in both communities in spite of the fact that some of his utterances of sentences containing the word "arthritis" are assigned conflicting truth-values in the two communities.

True: "arthritis" is a natural kind word and my last point exploits this fact. Burge might argue that the point would not carry over to "sofa", "contract", "square" or any other word with something like an analytic definition. But I (along with Bach 1988) think that this is quite inessential. If Alfred assumes mistakenly that rectangles are squares, I can ascribe to him, relative to our community, beliefs using the word "square". I can imagine doing so relative to a community in which "square" refers to squares and rectangles as well. Again the ascription of a belief to Alfred, relative to the counterfactual community, using the word "square" might be analyzed as an abbreviation for some longer utterance involving the words 'what they call "square" in Alfred's community'. But this is not to say that the ascription would be impossible. The point is that even though a word may have an analytic definition, to determine its contribution to the proposition expressed by an utterance of it, one may have to rely on contextual factors such as supplying in thought a mental counterpart of a suppressed expression referring to the particular use of the word in a community or to a person's beliefs.

IV.3. *What Is in a Community?*

Not only do I think parity is supported by pragmatic reasons, but there are reasons why, I think, we ought not to accept wholesale Burge's point about the role played by an individual's disposition to defer to other members of his speech community for the individuation of the contents of the individual's thoughts.

First, as emphasized in Burge (1986b, pp. 702–703), an individual's *disposition to defer* to experts from his community (and his readiness to accept corrections from them) may be — to some extent — crucial for the individuation of the individual's thoughts, say about arthritis:

The language does not present a standard of competence independent of individuals' activity. Minimal competence consists in conformity to the practice of others. "Greatest competence" consists in abilities to draw distinctions, to produce precisifications, to use numerous linguistic resources, to offer counterexamples to proposed equivalences — that elicit the reflective agreement of other competent speakers. We may imagine a vast, ragged network of interdependence, established by patterns of deference which lead back to people who would elicit the assent of others . . . incomplete understanding . . . does nothing to exempt me from responsibility to communal norms of evaluation when I have or express my . . . beliefs.

It is indisputable that non-expert speakers do defer to experts with respect to the application of specialized words such as "arthritis" (and why not "sofa"?): ordinary speakers just learn the meaning of specialized words by turning to the opinions of experts (say by consulting dictionaries or encyclopaedias). But does Burge think that the acquisition of any ordinary concept (such as the concept of a human thigh) involves the speaker's disposition to defer to members of his or her community? Suppose (presumably with Burge) that Alfred's disposition to defer to members of his community is crucial not just to the individuation of the content of his concept of arthritis but also to that of his concepts of thighs and joints. Imagine that in the counterfactual community the words "thigh" and "joint" are used nonstandardly so that it is assumed that there are joints in a person's thigh. Now relative to such a counterfactual community, Alfred might believe that he has arthritis in his thigh. Would Burge argue that we ought not to ascribe to him the belief that he has arthritis in his thigh but rather the belief that he has arthritis in his schmigh (where "schmigh" is an invented word in my language used to ascribe beliefs to members of the hypothetical community about what they call "thigh")? What, according to Burge, is a good candidate for being the concept of a thigh? And how to tell, by Burge's lights, whether the right ascription of a belief to Alfred, relative to the counterfactual community, ought to be "Alfred believes that he has tharthrititis in his thigh" or "Alfred believes that he has arthritis in his schmigh"? The content of Alfred's belief might be partly indeterminate — he might have standard concepts of thigh and joint and

a nonstandard concept of arthritis; or he might have non-standard concepts of thigh and joint and a conception of arthritis according to which it occurs only in joints (not in any tissue or muscle). We might, I claim, still use our words “thigh”, “joint” and “arthritis” to report, within our community, his fear that he has arthritis in his thigh.

Second, I ask, what about the disposition of expert members of a community to argue with experts of a different community? What should prevent us from “charitably” supposing that experts from the hypothetical community are disposed to open-mindedly examine the reasons why experts from the actual community do not take arthritis to be jointly a disease occurring in joints *and* muscles? Burge might correctly reply that there is no such thing as a community’s disposition to defer to the opinion of members of another community, or that such a disposition cannot meaningfully be predicated of a community. But the point is that individual experts from one community might be disposed to examine the opinions of experts from another community just as an individual layman might be disposed to defer to the opinion of an expert from his own community. Unless perhaps Burge assumes that a community has boundaries which its members cannot cross. But this view would have, it seems to me, no more plausibility than the view of the incommensurability between paradigms made notorious by Thomas Kuhn. Remember: nothing in Putnam’s Twin Earth thought-experiment entails such a view of the boundaries of a community, since it is agreed that expert chemists from both Earth and Twin Earth might converge on the view that “water” does not have the same extension on Earth and on Twin Earth.

V. AN INDIVIDUALIST ACCOUNT OF BURGE’S THOUGHT-EXPERIMENTS

I have expressed doubts about the joint acceptance of premisses (1) and (2) of the Burgian argument because I have doubts that “Alfred believes that he has arthritis” may be truly used to ascribe a belief to Alfred relative to the actual community *and* may not be so used to ascribe a belief to him relative to the counterfactual community. But suppose now we accept Burge’s conclusion *C* that what makes a belief-ascription true are not only facts internal to the person who is being ascribed the

belief but also facts pertaining to the person's social environment. Does it then follow that facts from the person's social environment are inevitably part of the contents of a person's thoughts? Burge thinks it does. As argued in Jacob (1987) along with Loar (1987a, b), I think there is room for denying that it does.⁹ So what is the further premiss used by Burge to derive the conclusion that the social environment enters the content of an individual's thought (or belief) from conclusion C, i.e. from the assumption that the social environment is part of the truth-conditions of the belief-ascription? I think the needed premiss is that the contribution made by the proposition expressed by the "that"-clause to the truth-conditions of the belief-ascription is identical to the content of the believer's thought. I want to provide reasons for rejecting this identification and therefore for rejecting the anti-individualist assertion that the environment is part of the contents of a person's thoughts. My reasons will also be reasons for not giving up content-dualism and reconcile it with due acknowledgment of the fact that the narrow content of a thought is not identical to the linguistic meaning of the sentence uttered to express it.

V.1. *An Utterance Is an Interpretation of a Thought*

Unlike a person's thought or belief which a belief-ascriber is trying to capture, and which is a mental representation, a belief-ascription is an utterance. Unlike the believer's mental state which purports to represent directly a certain state of affairs (say, that Alfred has arthritis in his thigh), the belief-ascription purports to represent directly the believer's mental state (Alfred's belief that he has arthritis in his thigh or his view of the condition of his thigh) and indirectly the state of affairs (the condition of Alfred's thigh). In a belief-ascription, the ascriber selects a sentence of his or her language which he or she embeds under the scope of the relevant psychological verb: the embedded sentence (or "that"-clause) is expected to provide an adequate or revealing characterization of the believer's thought about the state of affairs. What I am going to argue is that there are reasons for denying that the propositional contribution made by the "that"-clause to the whole proposition expressed by the belief-ascription (or to the belief-ascription's truth-conditions) is identical to the content of the believer's thought or belief.

Were the believer himself to express his thought or belief, he would utter a simple (unembedded sentence), such as “I now have arthritis in my thigh”. This sentence has a linguistic meaning, which (because of the presence of the indexical “I”) we can think of as a logical form with a free variable — not a complete proposition — with gaps to be filled by information drawn from the context of utterance. It expresses the proposition that Alfred has arthritis in his thigh. In virtue of its propositional content, this utterance is used by the speaker to express his thought. Both the thought expressed and the utterance used to express it can have a truth-value. If Alfred utters in our actual community “I now have arthritis in my thigh”, both his utterance and his thought are false and they have the same truth-conditions: they purport to represent or depict (mentally or linguistically) the same state of affairs. An advocate of the “theory of direct reference” would presumably assume that Alfred and Alfred’s thigh are parts of the “singular” (or Russellian) proposition expressed by Alfred’s utterance. Certainly Alfred and Alfred’s thigh are parts of the truth-conditions of Alfred’s utterance. Theorists of direct reference (or singular thought theorists) might assume that Alfred and Alfred’s thigh are *parts* of Alfred’s thought that he has arthritis in his thigh together with mental representations (or modes of presentation) of Alfred and Alfred’s thigh — say by assuming (as does Salmon 1986) that one constituent of the thought expressed is the ordered pair whose first coordinate is Alfred and whose second coordinate is a mode of presentation of Alfred or perhaps by assuming (as Evans and McDowell do) that a (*de re*) mode of presentation of an individual (or part of an individual) would not be “available” were the individual (presented) not to exist. So at least on one possible view of the proposition expressed by Alfred’s utterance, this singular (Russellian) proposition cannot be identical to Alfred’s thought since the former is a constituent of the latter.

What is relevant to the present argument is the question: given that Alfred’s thought and Alfred’s utterance have the same truth-condition, what is the relation between Alfred’s thought and Alfred’s utterance? My answer is that (as argued in Sperber & Wilson 1986, ch. 4), Alfred’s utterance is an *interpretation* of Alfred’s thought. By *interpretation* I mean a semantic relation different from the relation that holds between

on the one hand words or thoughts and on the other hand what satisfies them, what they refer to or what has to be the case for them to be true, viz. what makes them true or false. A linguistic or mental representation of a state of affairs is made true or false by the state of affairs (which it purports to describe). A fragment of a linguistic or mental representation can be said to refer to a constituent of a state of affairs. But unlike referential or truth-conditional semantic relations which hold between representations (whether mental or linguistic) and non-representations, an *interpretive* relation holds between two representations.

What makes a representation R a good interpretation of another representation R' is not that R is made true by R' . Consider the case of a translation: we have a set of utterances or inscriptions of sentences from one language used to stand for another set of utterances or inscriptions of sentences from another language. So translation is a relation between two sets of linguistic representations. What makes it the case that the former set can “stand for” the latter? Although there are correct and incorrect translations, we do not say that the translating utterances “are made true” by the translated utterances. We do not say that a piece of indirect quotation “is made true” by the reported utterances. Neither do we say, in the case of a summary, that the summarizing utterances “are made true” by the summarized utterances or that exegetical utterances “are made true” by the aphoristic utterances (subjected to exegesis). In the case of translation, a translator first grasps the set of propositions expressed by the sentences initially uttered in one language. He then looks for sentences in his or her language which will appropriately express the set of propositions in question. He will try to pick up in his language sentences with the same linguistic meaning (or semantic structure) as the sentences originally uttered.

In translation, what makes a translation a good one might be the identity between the set of propositions expressed by the translating utterances and the set of propositions expressed by the translated utterances. But what of a summary or an exegesis? By assumption, the two sets of propositions are not identical: the set of summarizing propositions is smaller than the set of summarized propositions and the set of exegetical propositions is larger than the set of initially expressed

propositions. So identity between two sets of propositions cannot be what makes a summary a good one any more than it can make an exegesis a good one. Rather the set of summarizing propositions must *resemble* the set of summarizing propositions, the former must be a good *prototype* or *sample* of the latter. Exegesis being the converse of summary, the set of initially expressed propositions must be a good prototype or sample of the exegetical propositions. In both cases, the two sets of propositions resemble each other — identity being a limiting case of similarity.

I want to apply my claim that interpretation holds by virtue of logical or propositional resemblance or similarity to the relation between a person's thought (or belief) and the utterance used to express it. Now there is a difference between on the one hand translation, summary and exegesis and on the other hand the expression of a thought by means of an utterance. In translation, summary, or exegesis the relation is between two sets of utterances. When an utterance expresses a thought the relation holds between an utterance and something else (a thought). Arguably, what makes a thought different from an utterance is that, unlike the sentence-type uttered which can be *quoted*, a thought-type cannot. However, granted that we do not know how to quote a thought, what makes it possible for utterances to be interpretively related to one another is that they express propositions. So if we assume that a thought may have a propositional form, structure and content, then a thought may be held to have all the required ingredients to enter into an interpretive relation with an utterance or with another thought. Hence there is no good reason not to view the interpretive relation between a thought and the utterance used to express it as based on the logical or propositional resemblance between them. I suggest that in virtue of its propositional form, an utterance can be held to *summarize* a thought.

Of course when I distinguish the interpretive relation between two representations from other semantical referential or truth-conditional relations, I am not denying that a representation can refer to another representation as when you say "This last proposition is intriguing" or "Paul's thought last Thursday was trivial". Then part of the utterance is supposed to refer to a proposition or a thought. But I am distinguishing interpretation from truth and reference.

V.2. *Content-Dualism Reconciled With Recognition That Narrow Content Is Not Linguistic Meaning*

I claimed at the end of the first section that content-dualism can accommodate the fact that the narrow content of a thought is *not identical* to the linguistic meaning of indexical expressions used to express it. Indeed to acknowledge that the linguistic meaning of “I” is not identical with the narrow content of the thought-constituent expressible by an utterance of “I” is compatible with the claim that the former is an *interpretation* of the latter, that they resemble each other. The speaker cannot offer his audience direct access to his first person thought-constituent about himself or herself. But in virtue of its linguistic meaning, “I” is the most straightforward route by which the speaker can direct his or her audience’s attention and thought-processes upon himself or herself (the speaker). What makes the linguistic meaning of “I” relevantly similar to the narrow content of a first person thought-constituent is that, for each utterance of a token of “I”, both are modes of presentation of the same person — they share the same reference on each occasion. As noted by McGinn (1983, ch. 5), there is the following parallel: on the one hand the linguistic meaning of “I” always presents its referent in the same way (as the speaker or the utterer of this token of “I”) but each token of “I” may refer to a different person. On the other hand although two persons may have different particular conceptions of themselves (different first person thought-tokens), each of those conceptions might be tokens of the same first person thought-type. And this might be true even though the linguistic meaning of the word-type “I” is not identical to the content of the first person thought-type expressed by any token of “I”.

In Evans’s example of the rational speaker expressing two thought-constituents with different narrow contents by means of two tokens of the same words (“this ship”) with one and the same linguistic meaning, assuming that the ship is descriptively identified, it is plausible to say that the linguistic meaning of “this ship” (perhaps “the ship most salient in the context of utterance”) is a *summary* or an abbreviation of the narrow content of the thought-constituent, *the ship visible to my left/right*, where, as we saw, summary is a typical interpretive relation that holds in virtue of logical resemblance. Perhaps the objection could be

made that I am presupposing that the rational speaker's mental mode of presentation of the ship is descriptive whereas actually the ship is presented (or identified) demonstratively (or perceptually), not descriptively. I take it that a third person with no perceptual access to the ship (such as the writer or the reader of the present paper) would however represent to himself or herself the rational speaker's demonstrative (or perceptual) mode of presentation of the ship descriptively — by means of something like *the ship visible to his [the rational speaker's] left/right*. Such a descriptive mode of presentation might be thought of as a second-order mental representation or interpretation (in the mind of a third person with no perceptual access to the ship) of the rational speaker's first-order demonstrative mode of presentation of the ship. The descriptive mode of presentation of the ship might therefore be thought of as an exegetical expansion of, or a conceptual comment upon, the demonstrative mode of presentation of the ship.

Now, as I said, the purpose of a (*de dicto*) belief-ascription (“Alfred believe that *p*”) is to characterize the believer's viewpoint. First, there are two aspects to the subject's state: his or her attitude and the content of his or her attitude. The former is described by the psychological verb preceding the “that”-clause. Second and more importantly, given that the belief-ascription has truth-conditions, there are two semantic relations to be distinguished: (a) the relation between either the proposition expressed by the ascriber's utterance of “*p*” (in the process of uttering the whole ascription) or the believer's thought and their common truth-conditions; (b) the relation between the proposition expressed by the utterance of “*p*” (in the process of uttering the whole sentence) and the content of the believer's thought.

V.3. *Two Mistaken Assumptions About a Thought Ascribed*

One of two assumptions might help the anti-individualist to the conclusion that the environmental facts which are part of the truth-conditions of the belief-ascriptions are parts of the content of the believer's thought:

(A) that the proposition expressed by the ascriber's utterance of “*p*” (and named by “that *p*”) is identical to the content of the believer's thought;

(B) that the proposition expressed by the ascriber's utterance of "*p*" is made true by the content of the believer's thought.

I reject both (A) and (B).

The belief-ascription as a whole can be said to describe the believer's state: if the believer is not in the state imputed to him by the ascription, the latter will be false. So belief-ascriptions have truth-conditions which may well, as emphasized by Burge, include facts from the believer's (social) environment. Arguably, the proposition expressed by the entire ascription may depend compositionally on the propositional contribution made by the "that"-clause. So "that *p*" may be taken to be a singular term naming or referring to the proposition expressed by the utterance of "*p*". But this does not mean that the proposition named by "that *p*" and expressed by the ascriber's utterance of "*p*" in the process of his whole utterance is identical to the content of the believer's thought.

A proposition or a thought may be referred to in a variety of ways, just like a number may be referred to in a variety of ways. But not all ways of referring to a proposition, a thought or a number supply a revealing characterization of the proposition, thought or number. I may refer to the number 6 as "my younger brother's favorite even number". This may or not enable you to pick out the number 6. If in response to questions such as "What did Paul say (or what does Paul think)?", I say: "Paul said (or thinks) what Mary said (or thought) last Friday", I may have successfully referred to a thought or a proposition expressed by an utterance by means of the relative clause "what Mary said (or thought) . . ." But I have not thereby automatically supplied a proposition such that if you grasp it you will be able to interpret the thought or proposition I have referred to. I can choose to name "Al" Paul's thought or the proposition he expressed, and refer to either as Al. Suppose François utters some French words. Suppose you do not understand French and you ask me "What did François say?". Suppose I faithfully reproduce another token of the French word-types François just uttered (I quote François's words). I have not thereby automatically provided for you (if you do not speak French) a revealing characterization of François's thought. Similarly, if I successfully refer to Paul's thought or to the proposition expressed by Paul's utterance by using the name "Al", I may have done so without expressing a proposition. It is one thing to

use a linguistic device to refer to a proposition or a thought; it is another thing to express a proposition capable of interpreting a thought.

In a belief-ascription, “that *p*” (just like “A1”) may well name, or refer to, the proposition expressed by an utterance of “*p*”. But it does not thereby name, or refer to, the believer’s thought. Rather, insofar as the job of the belief-ascription is to provide a revealing characterization of the content of the believer’s thought, the utterance of “*p*” expresses a proposition (referred to by “that *p*” or named “A1”) which is an *interpretation* — a *summary* — of the believer’s thought and which does its job (or fulfills its role) if it resembles appropriately the believer’s thought. But it need not be identical to it. In fact, if it is a summary of the believer’s thought, it cannot be identical to it. So I reject (A).

As I said, the belief-ascription as a whole has a truth-condition. It includes the “that”-clause as a component. When uttering “Alfred believes that *p*”, the belief-ascriber utters “that *p*” (as part of his whole speech act). Although he may thereby refer to the proposition expressed by his uttering “*p*” (a sub-part of his whole speech act), his utterance of “*p*” is not made true (or false) by the content of the believer’s thought. His utterance of “*p*” (in this context) purports to be faithful to (to adequately depict or resemble) the content of the believer’s thought in virtue of the proposition thereby expressed (and referred to by the larger constituent “that *p*”). So the overall belief-ascription may well be made true or false by a state of the believer’s (together with facts from the environment). But to assume, as (B) would have it, that the content of the believer’s thought makes true the proposition expressed by the ascriber’s utterance of “*p*” is to fall victim to a *level* confusion: the believer’s thought and the ascriber’s utterance of “*p*” purport to represent or describe the fact (a fact which is neither mental nor linguistic) that Alfred has or not arthritis in his thigh. In the process of expressing a proposition with the above truth-condition, the ascriber offers an interpretation of the content of the believer’s thought. Similarly, if I translate an utterance from French to English, the proposition expressed by my English utterance will have presumably the same truth-condition as the French utterance. But my English utterance (or the proposition I thereby express) is not made true by the French utterance. I therefore reject (B).

So the truth-conditions of the belief-ascription include as a com-

ponent an *interpretive* relation (not a truth-conditional one) between the content of the believer's thought and the proposition expressed by the ascriber's utterance of "*p*" (named by the utterance of "that *p*"). And the proposition expressed by uttering "*p*" during the utterance of the whole ascription is not itself made true by the content of the believer's thought. Rather both the utterance of "*p*" and the (content of the) believer's thought are made true or false by a state of affairs (say by the condition of Alfred's thigh).

I argued in section IV for what I referred to as the parity principle: grounds for using the word "arthritis" to ascribe to Alfred, relative to the actual community, the belief that he has arthritis in his thigh ought to be grounds for using the same word to also ascribe to Alfred, relative to the counterfactual community, the belief that he has arthritis in his thigh. Conversely, grounds for refusing to use the word "arthritis" to ascribe a belief to Alfred, relative to the counterfactual community, ought to be grounds for refusing to use the same word for ascribing a belief to Alfred relative to the actual community. The reason is that the concept expressed by the ascriber's use of the word "arthritis" must logically resemble the concept mentally represented in the believer's mind as part of his thought so as to preserve propositional resemblance between Alfred's thought and the proposition expressed by the ascriber's utterance of "*p*" (in the process of uttering the ascription). If resemblance is thereby achieved, relative to our community, it may be achieved, relative to the counterfactual community as well. If it cannot be achieved, relative to the counterfactual community, it can no more be achieved, relative to our community.

Consider the ascription to Alfred, relative to the actual community, by a medical authority of the belief that he (Alfred) has arthritis in his thigh: the medical expert utters "he has arthritis in his thigh" as part of his whole utterance (where by "he" the expert refers to Alfred and by "his thigh" he refers to Alfred's thigh). He thereby expresses a proposition which he takes to be false (a proposition made false by the condition of Alfred's thigh). Perhaps he refers to this false proposition by his utterance of "that he has arthritis in his thigh". However this false proposition is not identical to the content of Alfred's thought that he has arthritis in his thigh. On the assumption that they are identical, it becomes difficult (if not mysterious) to make sense of the fact that part

of Alfred's thought involves the mistaken assumption that arthritis can be located in one's thigh. However the proposition expressed by the expert's utterance of "he has arthritis in his thigh" (which does not involve the mistaken assumption in question) may be a good interpretation of Alfred's thought: in virtue of the similarity between the two, it may provide an adequate summary of Alfred's thought. If so, it seems to me, my own utterance of "he has arthritis in his thigh" could be held to summarize the thought Alfred has in the counterfactual community.

I am not saying that Burge would endorse the truth of (B), i.e. that the belief-ascriber's utterance of "*p*" (as part of his overall utterance) is made true (or false) by the content of the believer's thought or that he denies that the ascriber's utterance of "*p*" and the believer's thought are both made true (or false) by the same state of affairs. But I am suggesting that proper recognition of the fact that the proposition expressed by the ascriber's utterance of "*p*" is intended to interpret the content of the subject's thought by being an appropriate summary of it allows us to resist Burge's anti-individualist conclusion that what has to be the case for the ascriber's utterance of "*p*" to be true also constitutes the content of the subject's belief.

The view that the proposition expressed by the ascriber's utterance of a sentence containing say "arthritis" is an intended interpretation of the believer's thought expressible by the believer's use of "arthritis" makes room for the distinction (argued for e.g. in Bach 1988) between an individual's *concept* of arthritis and the meaning of the word "arthritis" as determined by expert members of the individual's community. Bach (1988, note 3, p. 88) sharply distinguishes between using "a word one incompletely understands" and thinking "with a concept that one incompletely understands". He accuses Burge of equating the former phenomenon, which he takes to be perfectly intelligible, with the latter, which he takes not to be so, on the ground that "to understand a concept" just is *to possess it*. Unlike Bach however, I am not so sure that thinking with a concept one incompletely understands does not make sense. If we merely assume that the meaning of a specialized scientific word is a (scientific) concept, then I think it makes sense to say of a non-scientist individual, who has an incorrect or partial grasp of the meaning of the word, that he thinks with a concept he incom-

pletely understands: his own concept would be a partial second order interpretive representation of the full scientific concept.

Just as it is possible to ascribe thoughts to an individual who is not a physicist using words from physics which he incompletely understands, one could — assuming concepts to be sets — characterize an individual's concept as his or her possession of a subset, or a superset of, the scientific concept. I would argue that just as there must exist a resemblance relation between the public meaning of the word “arthritis” and the individual's concept, the latter may resemble a scientific concept without being identical to it. Furthermore an individual layman may store a complex concept expressible by a word competently used by scientific experts as a child might entertain the metalinguistic representation of a word as a lexical entry referring to whatever phenomenon or magnitude the expert (or adult) takes it to refer to. The individual's concept might reduce to the lexical entry of the word expressing the concept, in which case we might be reluctant to ascribe the concept to the individual.

Let me sum up. In the present paper I have tried on the one hand to show that recognition that narrow content cannot be identical with linguistic meaning is not fatal to content-dualism. On the other hand I have tried to provide an individualist account of Burge's thought-experiments such that recognition that the truth-conditions of belief-ascriptions include aspects of the believer's (social) environment does not entail that those environmental aspects are thereby parts of the contents of the person's thoughts. This account is based on the view that the proposition contributed by the clause prefixed by “that” (in the belief-ascription) is an *interpretation* of the believer's thought.

NOTES

Thanks to Ned Block, Steven Davis, Lawrence Hirschfeld and François Recanati for discussion.

¹ As recognized by Tyler Burge (1986, p. 10), the main representative of social anti-individualism, “psychology is not a monolith. Different explanatory tasks and types of explanation coexist within it. In questioning the view that psychology is individualistic, I am not *thereby* doubting whether there are subparts of psychology that conform to the strictures of individualism”. I take the “subparts of psychology” (in the last sentence) to refer *inter alia* to the cognitive psychological study of “subdoxastic” mental states (in Stich's 1978b sense). And so Burge would acknowledge that an individual's subdoxastic states may be individualistically construed. As explicitly emphasized by Burge, his view applies to intentionally describable mental states. Where exactly does the boundary

between intentional and subdoxastic lie? This is question the anti-individualist ought to face.

² Fodor (1987, ch. 2) argues, I think convincingly, that the fact that my brain, unlike my doppelgänger's, contains H₂O molecules or lives in a body containing H₂O molecules is irrelevant to the individuation of our brain states.

³ Maxim (a) can be severely construed — as asserting the identity between mental content and truth-conditions — or liberally construed — as asserting that content should determine truth-conditions (perhaps as a function taking contexts as arguments and truth-conditions as values). Eliminativists with respect to content (e.g. Stich) typically construe (a) severely; dualists (such as Block and Fodor) construe it liberally.

⁴ I say "perhaps" for such a narrow content presupposes ("perhaps" wrongly) that Alfred's mode of identification of the ship is descriptive (non-demonstrative).

⁵ For purposes of refuting content-dualism, Baker (1987) imagines that on Mars there are no red objects, that Martians wear special glasses which are never removed and which are such that light reflected from grey objects affects Martians' transducers in exactly the way that light reflected from red objects affect Earthians' transducers on Earth. Consider a pair of doppelgängers, one on Earth, one on Mars. They have the same experience of red. But one's experience is caused by his perception of a red object; the other's experience is caused by his perception of a grey object. Were the Martian located on Earth and were he to remove his glasses, he would experience the perception of red objects as red. Were the Earthian located on Mars and were he to wear the special glasses, the content of his experience would be identical to the Martian's. Furthermore the inferential potential of the Martian's experience caused by his perception of grey objects is identical to the inferential potential of the Earthian's experience caused by his perception of red objects. The two mental states have different causes with many similar effects. Baker claims that her thought-experiment shows that no concept of narrow content can meet reasonable constraints on content — such as that "(mental) representations denote properties of distal objects to which their tokens are causally related" and that "if two tokens of (mental) representations denote different properties, then they are of different types of representation". By the first constraint, the Earthian's mental representation denotes red objects; the Martian's mental representation denotes grey objects. By the second constraint, both representations are distinct. I think that the second constraint begs the question against narrow content and should not be subscribed to by a content-dualist.

⁶ As noted by Steven Davis in a letter, on Burge's view, it is wrong to claim that, relative to the counterfactual community, Alfred believes (or thinks) that he has arthritis in his thigh upon waking up one morning when he does have this belief (or thought), relative to the actual community. Nor should we say that, in the counterfactual community, the doctor confirms Alfred's suspicion that he has arthritis in his thigh. Rather, relative to the counterfactual community, he believes that he has tharthritus in his thigh. However this is just to say that Burge jointly accepts what I call a few paragraphs below premisses (1) and (2) — a joint acceptance which can, I think, be questioned on grounds of what I call "parity" (see below), since "arthritis" might be the best word in my language to report or ascribe Alfred's state of mind, relative to the counterfactual community, even if "arthritus" is short for 'what they call "arthritis" in Alfred's community'.

⁷ The second difference has been emphasized by Keith Donnellan (1989).

⁸ It is an open question, I think, whether one ought to say in this case that the proposition that John is coming is a Grician *implicature indirectly* inferred by Edith from the proposition directly and misleadingly expressed by my utterance, i.e. that Edith's husband is coming. Or whether one ought to say that it is part of the propositional content *directly* expressed by my utterance by means of my use of an expression that Paul would or might have used. There are, I think, general reasons for assuming that

Grician pragmatic factors contribute to the content of the proposition explicitly (or directly) expressed by an utterance — not just to the content of the propositions indirectly conveyed by an utterance.

⁹ Loar (1987a) argues that “sameness of *de dicto* or oblique ascription” is neither a necessary nor a sufficient condition for “sameness of psychological content”. He produces (pp. 102–103) an elegant variation on Kripke’s (1979) puzzle about Pierre in which, when transported to London, unaware that “Londres” and “London” refer to the same city, Pierre (who, prior to his move to London, was disposed to assent to the French sentence “Londres est jolie”) is now disposed to assent to the English sentence “London is pretty”. So in both situations, we are led by ordinary principles of belief-ascription to attribute to Pierre the belief that London is pretty (a belief he would either express by uttering, or assenting to, a French sentence or by uttering, or assenting to, an English sentence). However, Loar argues convincingly that, on a *de dicto* oblique reading, the same ascription ought to be said on the two occasions to reveal distinct beliefs each with a different *psychological* content — with a different inferential potential.

REFERENCES

- Bach, K. (1982), “De re Belief and Methodological Solipsism”, in A. Woodfield (ed.), *Thought and Object*, Oxford: Clarendon Press.
- Bach, K. (1988), “Burge’s New Thought Experiment: Back to the Drawing Room”, *The Journal of Philosophy*, LXXXV, 2, 88–97.
- Baker, L. R. (1987), “Content by Courtesy”, *The Journal of Philosophy*, LXXXIV, 4, 197–213.
- Block, N. (1986), “Advertisement for a Semantics for Psychology”, in P. A. French, T. E. Uehling, Jr. & H. K. Wettstein (eds.), *Midwest Studies in Philosophy*, X, 1986, Minneapolis: U. of Minnesota Press.
- Burge, T. (1979), “Individualism and the Mental”, in P. A. French, T. E. Uehling, Jr. & H. K. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. IV, Minneapolis: U. of Minnesota Press.
- Burge, T. (1982), “Other Bodies”, in A. Woodfield (ed.), *Thought and Object, Essays on Intentionality*, Oxford: Clarendon Press.
- Burge, T. (1986a), “Individualism and Psychology”, *The Philosophical Review*, XCV, 1, 3–45.
- Burge, T. (1986b), “Intellectual Norms and Foundations of Mind”, *The Journal of Philosophy*, LXXXIII, 12, 697–720.
- Dennett, D. (1982), “Beyond Belief”, in A. Woodfield (ed.), *Thought and Object, Essays on Intentionality*, Oxford: Clarendon Press.
- Donnellan, K. S. (1966), “Reference and Definite Descriptions”, in S. P. Schwartz (ed.), *Naming, Necessity and Natural Kinds*, Ithaca: Cornell UP.
- Donnellan, K. S. (1989), “Belief Content and the External World”, paper delivered at the APA Pacific Division, in Berkeley, March 24, 1989.
- Evans, G. (1982), *The Varieties of Reference*, Oxford: Clarendon Press.
- Field, H. (1977), “Logic, Meaning and Conceptual Role”, *The Journal of Philosophy*, LXXIV, 379–409.
- Fodor, J. A. (1980), “Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology”, in *Representations, Philosophical Essays on the Foundations of Cognitive Science*, MIT Press.
- Fodor, J. A. (1987), *Psychosemantics, the Problem of Meaning in the Philosophy of Mind*, MIT Press.
- Harman, G. (1982), “Conceptual Role Semantics”, *Notre Dame Journal of Formal Logic*, 23, 2, 242–256.

- Jacob, P. (1987), "Thoughts and Belief-Ascriptions", *Mind and Language*, 2, 4, 301–325.
- Kaplan, D. (1977), "Demonstratives", mimeo, UCLA.
- Kripke, S. (1979), "A Puzzle about Belief", in A. Margalit (ed.), *Meaning and Use*, Dordrecht: Reidel.
- Loar, B. (1987a), "Social Content and Psychological Content", in D. Merrill and R. Grimm (eds.), *Contents of Thoughts*, Arizona University Press.
- Loar, B. (1987b), "Reply to Bilgrami: A New Kind of Content", in D. Merrill and R. Grimm (eds.), *Contents of Thoughts*, Arizona University Press.
- Loewer, B. & E. Lepore (1986), "Solipsistic Semantics", in P. A. French, T. E. Uehling, Jr. & H. K. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. X, Minneapolis: U. of Minnesota Press.
- McDowell, J. (1977), "On the Sense and Reference of a Proper Name", in M. Platts (ed.), *Reference, Truth and Reality*, London: Routledge & Kegan Paul.
- McGinn, C. (1982), "The Structure of Content", in A. Woodfield (ed.), *Thought and Object, Essays on Intentionality*, Oxford: Clarendon Press.
- McGinn, C. (1983), *The Subjective View, Secondary Qualities and Indexical Thoughts*, Oxford: Clarendon Press.
- Perry, J. (1977), "Frege on Demonstratives", *The Philosophical Review*, 86, 474–497.
- Perry, J. (1979), "The Essential Indexical", *Noûs*, 13, 3–21.
- Putnam, H. (1974), "The Meaning of 'Meaning'", in H. Putnam, *Philosophical Papers*, vol. II, Cambridge: CUP.
- Recanati, F. (1989), "Direct Reference, Meaning, and Thought", forthcoming in *Noûs*.
- Salmon, N. (1986), *Frege's Puzzle*, MIT Press.
- Sperber, D. & D. Wilson (1986), *Relevance, Communication and Cognition*, Cambridge, Mass.: Harvard University Press.
- Stich, S. (1978a), "Autonomous Psychology and the Belief-Desire Thesis", *The Monist*, 61, 4, 573–591.
- Stich, S. (1978b), "Beliefs and Subdoxastic States", *Philosophy of Science*, 45, 4, 499–518.
- Stich, S. (1983), *From Folk Psychology to Cognitive Science, the Case against Belief*, MIT Press.
- Stich, S. (1988), "Narrow Content Meets Fat Syntax", mimeo.
- Wettstein, H. K. (1986), "Does Semantics Rest on a Mistake?", *The Journal of Philosophy*, LXXXIII, 185–209.

*CREA, Ecole Polytechnique-CNRS,
1, rue Descartes,
75005 Paris,
France.*