SOCIAL EXTERNALISM AND CONCEPTUAL ERRORS

BY ÅSA MARIA WIKFORS

I. INTRODUCTION

Ever since Putnam and Burge launched their respective attacks on individualist accounts of meaning, the individualist has felt squeezed for space.1 Very little room for manoeuvre, it seems, is left for philosophers who want to deny that meaning and mental content depend on the speaker’s social environment. One option, popular amongst individualists, is to grant that reference is socially determined, but argue that there is nevertheless a notion of meaning or content that can be understood individualistically. That is, the individualist can opt for a ‘narrow’ way of construing thought-content according to which thought-content is independent of reference and individuated purely internalistically.2 The problem with this option is that it entails an admission of substantial defeat. By construing thought-content narrowly, the individualist gives up on reference and is forced to cut the traditional link between thought-content and truth-conditions.

The only obvious alternative for the individualist is to deny that reference is socially determined, and to reject the conclusions of the anti-individualist thought-experiments. For instance, the individualist can deny that Burge’s patient uttering ‘I have arthritis in my thigh’ is to be interpreted standardly,  

as speaking of arthritis, and suggest, instead, that the patient has his own ‘arthritis’-concept, ‘arthritis’ (a disease of the joints and the ligaments), and has therefore expressed a true belief. This would leave individualism intact, with reference and content alike determined individualistically. But the trouble with this option is that it seems so thoroughly unattractive. Are we really to say that Burge’s patient has his own ‘arthritis’-concept, and expresses a true belief when uttering ‘I have arthritis in my thigh’? Would that not imply such a fragmentation of concepts and reference across the community of speakers that no common core of thoughts would remain? If every time individuals depart from the community practice they are to be ascribed their own concept, it appears that no two individuals in the community ever share the same concepts.

Individualists therefore seem faced with a dilemma: they can give up on reference, thereby compromising their individualism; or they can reject the communitarian conclusions and accept conceptual and referential fragmentation. Faced with these options, even strong-headed individualists will be tempted to desert to the communitarian camp.

In this paper I try to instil some courage into individualists by urging that we should rethink the apparent dilemma. We have grown accustomed to thinking of the anti-individualist thought-experiments in a particular way, which, I argue, is not unavoidable but rests on a problematic philosophical assumption. Once this assumption is made explicit, it is clear that the individualist can reject the communitarian conclusions without having to accept conceptual and referential fragmentation. I shall focus on Burge’s well known thought-experiment concerning the concept of arthritis, but the discussion applies generally to the whole range of thought-experiments designed to show that meaning and mental content are determined by the individual’s social environment. I shall not, however, be concerned with the threat posed to individualism by ‘physical externalism’, the thesis advanced by Putnam, Kripke and others that the meaning of certain types of terms (most prominently, natural-kind terms) is individuated by the speaker’s physical environment. Although I believe that related objections apply to this brand of externalism, I shall not attempt to show that here. As Burge emphasizes (p. 79), social externalism can be applied to all kinds of terms (‘sofa’, ‘brisket’, ‘red’, etc.), and not just to natural-kind terms: ‘The argument has

---


an extremely wide application. It does not depend, for example, on the kind
of word “arthritis” is. We could have used an artefact term, an ordinary
natural-kind word, a colour adjective, a social role term, a term for a
historical style, an abstract noun, an action verb, a physical movement verb,
or any of various sorts of words.’ So it is clear that the more radical threat to
individualism is posed by social externalism, and I shall limit my discussion
to that threat.

The core of the paper is divided into three sections. In §II I shall give a
brief account of Burge’s famous thought-experiment and the central notion
of incomplete understanding. In §III I shall spell out the problematic as-
sumption underlying Burge’s thought-experiment, and argue that this
assumption should be rejected. The final section draws out the consequences
for individualism of rejecting this assumption, and provides individualists
with a general recipe for how to withstand the communitarian thought-
experiments.

II. INCOMPLETE UNDERSTANDING

A brief summary of Burge’s thought-experiment will suffice. The argument
proceeds in three steps. First, Burge (p. 77) hypothesizes a speaker who is
‘generally competent in English, rational and intelligent’. This speaker has a
large number of attitudes commonly attributed with content-clauses con-
taining ‘arthritis’, including a number of true beliefs. However, the speaker
also believes falsely that he has developed arthritis in his thigh. This belief is
false, Burge argues, since though the speaker does not know this, the con-
ventions of his community are such that ‘arthritis’ does not apply to ailments
outside the joints. Secondly, Burge considers a counterfactual world where
all is as before except that in the counterfactual speech community ‘arthritis’
applies not only to a disease of the joints but also to various other rheuma-
toid ailments. This means, according to Burge, that in the counterfactual
world ‘arthritis’ expresses a different concept (the concept of a disease of the
ligaments as well as of the joints). Burge then concludes in the final step that
a member of the counterfactual community, in all ‘internal’ aspects identical
to the actual speaker, who utters ‘I have arthritis in my thigh’, is not mis-
using the word ‘arthritis’, but using it correctly. He is expressing a true belief.
As a result, meaning and mental content vary with social environment.

A central component of Burge’s reasoning is the notion of a non-
empirical or conceptual error. The speaker uttering ‘I have arthritis in my
thigh’, Burge says (p. 82), does not make an ‘ordinary empirical mistake’,
but a conceptual one. The speaker has misunderstood the concept of arthritis, yet
he has beliefs containing the concept. This, Burge argues, is an essential pre-
supposition of the thought-experiment. If the thought-experiment is to work,
he says (p. 83), ‘one must at some stage find the subject believing (or having
some attitude characterized by) a content, despite an incomplete under-
standing or misapplication. An ordinary empirical error appears not to be
sufficient.’

It should be clear why this is so. If Burge’s arthritis-patient in uttering
‘I have arthritis in my thigh’ does not make a conceptual error, and if
‘Arthritis affects the joints only’ is an empirical claim, then the fact that the
counterfactual community rejects this claim would imply not that their word
‘arthritis’ must have a different meaning, but that the disagreement between
the two communities is one of theory. That is, the conclusion to be drawn
would be that the counterfactual community has developed a slightly differ-
ent theory about the same disease, not that the counterfactual community
has a different ‘arthritis’-concept and speaks of a different disease.5 There
would then be no reason to say that the speaker in the counterfactual world,
when uttering ‘I have arthritis in my thigh’, expresses a belief different from
that of the speaker in the actual world.

The thought-experiment therefore relies on two central assumptions, as
Burge himself makes quite clear; first, that the individual makes a concep-
tual rather than an empirical mistake, i.e., that ‘Arthritis affects the joints
only’ plays such a special role within our English-speaking community that a
counterfactual community which rejects this statement must be one where
‘arthritis’ has a different meaning and expresses a different concept; and
secondly, that conceptual errors of this sort do not prevent standard belief-
attribution. The individual who utters ‘I have arthritis in my thigh’ therefore
makes a conceptual error, but can none the less be ascribed the standard
concept of arthritis, and thus can be said to hold the belief that he has
arthritis in his thigh.

Burge’s arguments for the second assumption are familiar. The reason
why the patient should be ascribed the community concept, despite his in-
complete understanding of it, is that he is committed to the community
practice and intends to speak of the same disease as the medical experts
within his community do. This commitment is shown by the fact that the
patient, when discussing his illness with his doctor, is willing to stand corrected:
‘The subject’s willingness to submit his statement and belief to the
arbitration of an authority suggests willingness to have his words taken in
the normal way – regardless of mistaken associations with the word’ (p. 101).
However, Burge does not hold that the speaker’s commitment alone implies

(1998), pp. 470–82, at pp. 478–9, for a related point.
that he should be ascribed the standard concept. The individual must also, Burge argues, ‘maintain a minimal internal and rational coherence and a broad similarity to others’ use of the language’ (p. 114). Thus Burge does not take the extreme view that the individual’s words are always to be interpreted standardly, regardless of how he uses them. If interpreting an individual standardly implies violating the requirements of rationality and coherence, Burge argues, we should reinterpret accordingly, despite the fact that the speaker is committed to the community practice: ‘When the deviance is huge attributions demand reinterpretation of the subject’s words’ (p. 91).

The upshot is that Burge’s anti-individualism relies on the idea that there is a certain category of conceptual errors that speakers commonly make. These are errors that are not so radical as to force a reinterpretation of the concepts associated with the speaker’s words and to prevent standard belief-attribution. That is, the errors are not such that they warrant the conclusion that the speaker simply has a non-standard concept. But at the same time the errors must be sufficiently radical to be genuinely conceptual, or else the speaker cannot be said to have an ‘incomplete grasp’ of the concept in question, and the anti-individualist conclusions will not follow. This is the idea that needs to be examined. I shall take a closer look at Burge’s notion of a conceptual error.

III. CONCEPTUAL ERRORS

Burge’s talk of conceptual errors and conceptual truths seems to suggest that he rejects Quine’s criticisms of the analytic/synthetic distinction. Nevertheless Burge expressly claims to have taken Quine’s critique on board. That the subject makes a conceptual error, Burge says (p. 88), should not be taken to suggest that the error is ‘purely’ a mistake about concepts, involving no error about the empirical world: ‘With Quine, I find such talk of purity and mixture devoid of illumination or explanatory power’. And yet, even though there are no such things as purely conceptual truths, Burge argues, there are ‘analytically’ true and ‘analytically’ false beliefs that are linguistic in the sense that ‘they are tested by consulting a dictionary or native linguistic intuitions, rather than by ordinary empirical investigation’ (p. 100). Thus, Burge says (p. 78), while the line between conceptual and empirical error is ‘pretty fuzzy’, in the case of the speaker uttering ‘I have arthritis in my thigh’ we must say that his error is conceptual in nature, since in the actual community ‘arthritis’ does not apply to ailments outside the joints ‘by a standard, non-technical dictionary definition’.

© The Editors of The Philosophical Quarterly, 2001
Burge’s suggestion, therefore, is that we can endorse what might be called a ‘weak’ notion of analytic or conceptual truths, one which is quite compatible with Quine’s rejection of the analytic/synthetic distinction. ‘Arthritis afflicts the joints only’ is a conceptual truth in the sense that this is what the dictionary tells us. To say that the speaker uttering ‘I have arthritis in my thigh’ makes a conceptual error is just to say that he goes against the dictionary definition. It does not require a commitment to a strong notion of conceptual truths.

This is clearly unsatisfactory. One might of course stipulate that whatever is found in the dictionary is ‘weakly analytic’ (i.e., analytic in a sense which does not invoke any principled distinction between the empirical and the non-empirical), but this notion of analyticity cannot do the work Burge wants it to do. While the dictionary may define ‘arthritis’ as ‘a disease of the joints only’, we cannot conclude from this that ‘arthritis’ must have a different meaning in the counterfactual community where the experts hold that ‘arthritis’ applies to a disease outside the joints, since the broader use in the counterfactual community may be attributable simply to a difference in theory, not in meaning. To draw the conclusion that the difference in expert use signals a difference in meaning, rather than a difference in theory, presupposes a strong notion of conceptual truth, one which clearly is not compatible with Quinean scepticism about the analytic/synthetic distinction.6

Burge’s thought-experiment therefore presupposes that ‘Arthritis afflicts the joints only’ is non-empirical in the strong sense of being an analytic or conceptual truth. And the question is why we should accept this assumption. Why should we accept the claim that ‘arthritis’ must have a different meaning in the counterfactual community? It may be that the belief that arthritis afflicts the joints only is central to our understanding of arthritis, but what gives Burge the confidence to say that it is so central that giving it up must imply a change in the meaning of the term ‘arthritis’? After all, medical terms like ‘arthritis’ play a complex role in medical theory, and as always with such terms, it seems possible to have a change in certain parts of the theory, including central parts, without any change in meaning. In this respect, ‘arthritis’ differs from what Putnam famously called one-criterion words, such as ‘bachelor’ or ‘vixen’ – words that are not caught up with theory but could plausibly be given simple definitions.7

6 That there is a tension in Burge’s views on this issue is noted by Sören Haggevist, Thought-Experiments in Philosophy (Stockholm: Almqvist & Wiksell, 1996). In opposition to Burge’s claim that he makes no use of any positive notion of synonymy in his thought-experiments, Haggevist observes (p. 176) that Burge’s thought-experiments ‘do presuppose some notion of synonymy, however unspecific it is left in Burge’s exposition’.

In fact in his more recent writings Burge himself puts forward considerations that support this line of reasoning. Instead of speaking of conceptual or conventional truths, Burge now speaks of ‘meaning-giving characterizations’, characterizations that ‘provide linguistic meaning – set a norm for conventional understanding’. Examples of meaning-giving characterizations, Burge suggests, are ‘A knife is an instrument consisting of a thin blade with an edge for cutting, fastened to a handle’, and ‘To walk is to move on foot at a natural unhurried gait’. A characterization is meaning-giving if the most competent speakers have reached equilibrium on it. However, Burge argues, meaning-giving characterizations differ from analytic truths in two important respects. First, their truth depends not merely on the meaning of the words involved, but also on the empirical world. Stating a meaning-giving characterization is not stating a ‘degenerate truth’, but stating empirical facts. Secondly, and relatedly, meaning-giving characterizations are dubitable and can be incorrect, even when there is complete agreement among the most competent speakers on these characterizations (p. 706).

Once we reject the idea that there are statements true in virtue of meaning alone, Burge argues, we have to recognize the mix of empirical and linguistic elements in all of our statements, and this mix shows that meaning-giving characterizations are dubitable even by the experts. The appeal to meaning-giving characterizations, according to Burge, is therefore perfectly compatible with Quine’s holistic picture of belief revision. Even our most central beliefs, those we take to be definitions, can be revised without necessarily implying a change in meaning: ‘Holism notes that the definitions that capture the conditions that the speaker treats in his usage as most fundamental for applying the term may be false. This shows, contrary to what some traditionalists presumed, that the definition cannot exhaust the significance of the term or the associated concept.’

If one takes this holistic picture seriously, as Burge himself clearly does, the original thought-experiment must be seen in a different light. On this view, ‘Arthritis afflicts the joints only’ plays the role of a meaning-giving characterization – one on which the experts all agree and which is taken to be central to the meaning of ‘arthritis’. However, since meaning-giving characterizations are revisable, it can no longer be concluded that the speaker who does not accept this characterization has an incomplete understanding of the concept of arthritis and makes a conceptual error when

9 Burge, ‘Concepts, Definitions, and Meaning’, Metaphilosophy, 24 (1993), pp. 309–27, at p. 316. See also p. 314: ‘scientific practice indicates that a definition that functions as the most basic explanation of a concept at one time can later be displaced’.
uttering ‘I have arthritis in my thigh’. The experts would take him to have made a mistake, of course, but the experts themselves may be wrong. Consequently it cannot be concluded that ‘arthritis’ must express a different concept in the counterfactual community. The conclusion to be drawn is not that the counterfactual community has a different ‘arthritis’-concept, but that one of the communities (ours or theirs) is mistaken in its ‘meaning-giving characterizations’.

There are two possible responses to this. The first is simply to say that whether ‘Arthritis afflicts the joints only’ is or is not a conceptual truth in English is beside the point. Instead, we should grant Burge that ‘Arthritis afflicts the joints only’ is a conceptual truth, and then proceed to meet his arguments. After all, Burge is not doing descriptive linguistics but exploring a philosophical thought-experiment. However, as is so often the case in thought-experiments, caution is called for. Burge explicitly makes a claim about English (the actual community is said to be our normal English-speaking community) and much of his case appeals to our intuitions about the English language. If Burge were simply stipulating that, in a possible language, ‘Arthritis afflicts the joints only’ expresses a conceptual truth, then he could not rest his argumentation, as he now does, on our actual linguistic practices.

The second reply consists in questioning the idea, accepted by Burge himself, that the thought-experiment presupposes that ‘Arthritis afflicts the joints only’ is a conceptual truth, and that the patient in the actual world makes a conceptual error when asserting ‘I have arthritis in my thigh’. All that is needed for the thought-experiment to go through, it might be argued, is the assumption that the patient’s utterance ‘I have arthritis in my thigh’ is false in the actual world and true in the counterfactual world. That is, it is because there is a difference in the extension of ‘arthritis’ in the two worlds that it follows that there must be a corresponding difference in concepts. Consequently we can get the conclusion that there is a difference in thought-content without relying on problematic assumptions about conceptual truths.

It is, however, difficult to see how this reply helps. The question is: what is it that on the social externalist view grounds the alleged difference in extension? The answer, it seems, is that the difference in extension derives from the difference in community conventions and community definitions. Burge is quite explicit about this. The only difference between the actual and the counterfactual community, he emphasizes, is that in the actual community the convention is such that ‘arthritis’ applies to a disease of the joints only, whereas in the counterfactual community ‘arthritis’ has a wider application. Far from removing the need to appeal to conventional or...
conceptual truths, the difference in extension itself relies on there being such truths.\textsuperscript{10} By questioning the idea that there is a difference in conventions between the two communities, we therefore also question the idea that there is a difference in extension. It can no longer be asserted that what the patient in the actual world has in his thigh is not in the extension of ‘arthritis’. Indeed, the counterfactual community could as well be our future community, one in which medical discoveries have made the experts conclude that arthritis can affect the ligaments as well as the joints. (To borrow one of Putnam’s examples, it could be that in the second community medical science is more advanced and it has been discovered that the inflammation of the joints called ‘arthritis’ is caused by a virus which in fact also causes inflammation of the ligaments.) After all, if the claim ‘Arthritis affects the joints only’ really is an empirical statement, rather than a conceptual truth, then it is not up to us to legislate its truth-value.

There is of course a type of externalist argument that does not seem to rely on the notion of conceptual or definitional truth, i.e., the argument employed by Putnam, Kripke and others in the case of natural-kind terms. Putnam and Kripke explicitly deny that it is a truth of meaning that water is H\textsubscript{2}O. And yet, they say, if water is indeed H\textsubscript{2}O there is no possible world in which water is not H\textsubscript{2}O. Thus a counterfactual world in which ‘water’ does not refer to H\textsubscript{2}O is a world in which ‘water’ means something different from what it means in the actual world (assuming that ‘water’ refers to H\textsubscript{2}O in the actual world).\textsuperscript{11} We can therefore get the result that there is a difference in meaning just by appealing to the idea that there is a difference in extension, without relying on anything like the analytic/synthetic distinction.

Assuming, for the sake of argument, that this is correct and that physical externalism can get by without appealing to conceptual truths,\textsuperscript{12} the question is how considerations of this sort can be used to support social externalism. According to Putnam, what grounds the idea that there is a difference in the extension of ‘water’ in the two worlds is not that there is a difference in expert theory, but that there is a difference in the physical environment, i.e., in the stuff picked out as ‘water’ in the two worlds. (Putnam

\textsuperscript{10} ‘Individualism and the Mental’, p. 78. In his ‘Other Bodies’, in Woodfield (ed.), Thought and Object, pp. 97–120, Burge even tries to derive the difference in extension of ‘water’ in the twin earth case from a difference in conventions: ‘The fact that the twin earthians apply “water” to XYZ is not a reflection of a shift in extension of an indexical expression with a fixed linguistic (English) meaning, but of a shift in meaning between one language, and linguistic community, and another. Any expression, indexical or not, can undergo such “shifts”, as a mere consequence of the conventionality of language’ (p. 105).


\textsuperscript{12} For some scepticism on this point see Bob Hale, ‘Modality’, in B. Hale and C. Wright (eds), A Companion to the Philosophy of Language (Oxford: Blackwell, 1997), pp. 487–514.
claims that in 1750, when there was no difference in expert theory between earth and twin earth, ‘water’ none the less had a different extension on the two planets.) This is why arguments concerning natural-kind terms do not have to rely on appeals to conventions or definitional truths. But this strategy is not open to social externalists, since again their case relies solely on social differences between the two worlds. Of course one might decide to endorse physical externalism instead of social externalism, but the problem then is how to extend the externalist case beyond that of natural-kind terms. It seems difficult to do so in the case of ‘arthritis’, since diseases are notoriously bad candidates for natural kinds, and the other words Burge discusses (‘sofa’, ‘red’, ‘brisket’, ‘contract’) are even poorer candidates for a physical externalist treatment.

In order to support the conclusion that there is a difference in concepts and thought-content between the two worlds, therefore, Burge must rely on the assumption that the speaker in the actual world makes a conceptual error, that he has an incomplete understanding of the concept of arthritis. But this assumption, as argued above, is very questionable, and it is one about which Burge himself, given his commitment to holism, should be sceptical. If this is right, the next question is what follows for individualism.

IV. INDIVIDUALISM RECONSIDERED

Burge’s thought-experiment, I suggested above, seems to present individualists with a dilemma: internalization of content, or fragmentation of reference and concepts. It should now be clear that the dilemma can be escaped by questioning the assumption underlying the thought-experiment, that the speaker makes a conceptual error. Once this assumption is questioned, individualists can grant Burge the first step of the thought-experiment, thereby avoiding referential and conceptual fragmentation, and yet resist the anti-individualist conclusions. That is, they can grant the claim that the patient has the standard concept of *arthritis*, rather than his own concept, and that he (most likely) makes an error when uttering ‘I have arthritis in my thigh’; but they can point out that this does nothing to support social externalism. Since the error in question is not conceptual, it can no longer be claimed that the patient’s twin in the counterfactual world has a different concept, and so nothing follows about the social determination of concepts. What justifies ascribing the standard concept to the patient is not his reliance on the community experts but the fact that by and large he uses the word ‘arthritis’ in accordance with standard use. (Burge describes the speaker as making a wide range of correct ‘arthritis’-judgements: see
‘Individualism and the Mental’, p. 77.) Of course, as emphasized above, once it is denied that ‘Arthritis afflicts the joints only’ expresses a conceptual truth, it is an open question whether in fact the patient makes an error. However, the point is that individualists can allow for the possibility that the patient makes an error without thereby compromising their individualism.

Instead of taking this line, individualists have been far too willing to grant the set-up of the anti-individualist thought-experiments. A good illustration of this can be found in Davidson’s discussion of Burge. Davidson criticizes Burge, and rejects the idea that, as Davidson puts it (‘Knowing One’s Own Mind’, p. 448), ‘what we mean and think is determined by the linguistic habits of those around us’. To make his point Davidson considers a version of Burge’s arthritis example. Suppose, Davidson says, that he, who believes the word ‘arthritis’ applies to inflammation of the joints only if caused by calcium deposits, and Arthur, who knows better, both sincerely utter ‘Carl has arthritis’. According to Burge, Davidson says, the two speakers must then mean the same thing by their words, and we should report them as having the same belief. But, Davidson argues, it is not correct that his and Arthur’s beliefs should be reported in the same way. If Smith were to report that Davidson and Arthur both believe that Carl has arthritis, he would be misleading his audience unless he added ‘But Davidson thinks that arthritis must be caused by calcium deposits only’. This shows, Davidson argues (p. 449), ‘that the simple attribution was not quite right; there was a relevant difference in the thoughts Arthur and I expressed when we said “Carl has arthritis”’.

What Davidson suggests, therefore, is that his believing that arthritis is caused by calcium deposits only implies that his ‘arthritis’-thoughts are not shared with Arthur. But this, clearly, is a problematic suggestion. If whenever two speakers disagree about the causes of arthritis it follows that they must have ‘different thoughts’, then it is hard to see how two speakers could ever share any thoughts at all. There are countless disagreements of this sort within a community, and if every such disagreement implied a conceptual disagreement, there could be no sharing of concepts, and no such thing as disagreement in belief. This is why it is essential that individualists should not uncritically accept the set-up of the anti-individualist thought-experiment, i.e., the assumption that the speaker in the experiment has an incomplete understanding of the relevant concept. As long as individualists accept this assumption, as for example Davidson does, the dilemma sketched above will seem unavoidable.

To this it might be objected that all that has been shown so far is that Burge has picked a bad example. What Burge needs to support the anti-individualist conclusions, it could be argued, is a case where it is
uncontroversial that the speaker makes a conceptual error – for instance, a case involving the kind of one-criterion words that Putnam discussed, such as ‘bachelor’ or ‘vixen’. The individualist strategy suggested above could not then be employed, since there is no doubt that the error in question is conceptual, and the anti-individualist conclusions would follow.

Interestingly, Burge himself discusses such a case in his ‘Belief and Synonymy’ (Journal of Philosophy, 75 (1978), pp. 119–38), which predates the famous thought-experiment. He takes his starting-point in the well known 1950s debate sparked by Benson Mates’ suggestion that two terms may be synonymous and yet not interchangeable in belief contexts. For instance, someone might doubt ‘Whoever believes that all Greeks are Greeks believes that all Greeks are Hellenes’, but not ‘Whoever believes that all Greeks are Greeks believes that all Greeks are Greeks’. The problem presented by Mates’ argument, Burge says, is that it seems to force defenders of traditional accounts of substitution in belief contexts (i.e., defenders of the idea that synonyms are logically interchangeable in belief contexts) either to hold that a person entertaining such doubts believes analytical falsehoods, or to argue that such a person is not speaking English. But, Burge argues (p. 124), neither option is plausible: it is unacceptable to attribute an analytical falsehood to a speaker when he shows ‘no sign of exceptional stupidity or irrationality’, and it is utterly implausible to say that the speaker is not competent in English. Instead, Burge offers a third option.

To illustrate this third option, he considers the following statements: ‘I once believed that some female foxes are not vixens, since I thought that a fox was not a vixen unless it had had sexual intercourse, but I never thought that some female foxes are not female foxes’, and ‘For years I believed that a fortnight was ten days, not fourteen, though of course I never believed that fourteen days were ten days’. The terms ‘fortnight’ and ‘fourteen days’, Burge says, are synonyms, and so are the terms ‘female fox’ and ‘vixen’. Even so, he argues, these statements about past beliefs are not nonsensical, but should be taken at face value as being possibly true. A speaker can be said to entertain the belief that some female foxes are not vixens, or that a fortnight is ten days, without thereby being irrational. This is so because it is quite possible to show that the speaker’s mistaken belief is rational by appealing to the possibility that he is linguistically mistaken. The speaker’s false belief stems from linguistic misinformation, and this, Burge argues (p. 129), ‘deprives one of grounds for holding that such a belief would be irrational’.

Burge’s suggestion, therefore, is that by appealing to the fact that the speaker is linguistically misinformed, the conclusion that the speaker is not speaking English can be avoided without thereby describing him as being, to
use Burge’s words, ‘exceptionally stupid and irrational’. It is not irrational to believe that a fortnight is ten days if one has misunderstood which concept is expressed by the word ‘fortnight’. If Burge is right about this, then he appears to have an argument leading to the conclusion that the community norms are constitutive of meaning and mental content. If speakers who assent to ‘A fortnight is not fourteen days’ can be said to believe that a fortnight is not fourteen days because of their commitment to the community practice, then it would be easy to design a thought-experiment showing meaning and mental content to be determined by the ‘social environment’ (in a counterfactual community where ‘fortnight’ means ten days, the speaker assenting to ‘A fortnight is not fourteen days’ would then have made a true statement and expressed a true belief). Indeed, Burge already draws these anti-individualist conclusions in ‘Belief and Synonymy’, pp. 134–5: ‘Communal conventions about the meaning of a speaker’s words tend to override what a speaker mistakenly associates with his words in determining what he says and even, sometimes, believes’.

Does this suggestion work? In what sense could appealing to the notion of linguistic misinformation avoid an unacceptable violation of the rationality requirement? Well, if we say of someone that he believes that a fortnight is ten days, but add that he does not know that ‘fortnight’ means fourteen days, then it hardly seems as if we have described the speaker as holding an irrational belief. But the reason for this, clearly, is that by adding that the individual in question is linguistically misinformed, we make clear that he uses ‘fortnight’ non-standardly and cannot be attributed the concept of fortnight. That is, we make clear that the individual does not believe that a fortnight is ten days. Burge, in fact, acknowledges that there is a sense in which this is so. In a sense, Burge says, ‘the speaker took “fortnight” to mean “period of ten days” when he used it’ (‘Belief and Synonymy’, p. 130). However, Burge argues (p. 131), to conclude from this that the speaker’s words should be reinterpreted would be to ignore an important aspect of the speaker’s attitude: ‘The willingness of the speaker to submit his statement to the arbitration of a dictionary indicates his commitment to having his words taken in their conventional sense, whatever that sense is’. This is familiar from ‘Individualism and the Mental’, but it is clearly unsatisfactory as an answer to the question under consideration. Burge’s question was how we could interpret the speaker as believing in an analytic falsehood without describing

---

13 This suggestion can be found also in Burge, ‘Individualism and the Mental’, p. 100. Responding to the objection that charity requires reinterpreting speakers who make conceptual errors, Burge says ‘There is nothing irrational or stupid about the linguistic or conceptual errors we attribute to our subjects. The errors are perfectly understandable as results of linguistic misinformation.’
him as being irrational. And it is clear that an appeal to his commitment to
the shared rules cannot help with this question; if this commitment really
implies that ‘fortnight’ expresses the standard concept, then it follows that
the speaker is to be attributed the belief that fourteen days is ten days, which
is to say that it follows that the speaker is to be described as being
‘exceptionally stupid and irrational’.14

For this reason I think it is no accident that Burge in ‘Individualism and
the Mental’ has moved away from this type of example. The ideas that
drive the thought-experiment involving the concept of arthritis are simply
missing when it comes to clear cases of conceptual truths, such as ‘A vixen is
a female fox’ and ‘A fortnight is fourteen days’. Whereas in the ‘arthritis’
case we are strongly inclined to say that the patient does not have his own
‘arthritis’-concept, but rather makes an error, this is not true in the case of
the speaker asserting ‘A fortnight is ten days’. A speaker who makes that
assertion is not a speaker who mistakenly believes that a fortnight is ten
days, but a speaker who lacks the concept of fortnight and uses ‘fortnight’ to
mean a period of ten days. Of course, as Burge suggests, the speaker makes a
mistake of a sort – a linguistic mistake. That is, he is mistaken about the
meaning of ‘fortnight’ in English, and he would most likely stand corrected
when told of his mistake. However, a linguistic mistake is not a conceptual
mistake, and it can hardly entail anti-individualism. To support those con-
clusions it has to be argued that the speaker has beliefs containing the
concept of a fortnight, and that, again, is what has not been shown.

This provides us with a general diagnosis of the anti-individualist
thought-experiments. What gives these experiments their apparent force is
the use of examples that do not involve obvious conceptual truths, but are
border-line cases like ‘Arthritis afflicts the joints only’ or ‘Arthritis is caused
by calcium deposits only’. What characterizes these statements is that they
are central to our understanding of the concept in question, but yet are not
straightforwardly conceptual truths like ‘A vixen is a female fox’ or ‘A
fortnight is fourteen days’. Anti-individualists can then exploit the border-
line quality of these statements. That is, they can exploit the fact that we are
not inclined to say that the speaker who rejects such a statement has a non-
standard concept, as we are in the case of an individual rejecting conceptual
truths. However, as I have argued above, the very fact that these cases do
not involve clear instances of conceptual truths prevents the thought-
experiment from going through – it prevents the conclusion that the
difference between the two communities is conceptual.

14 See Bilgrami’s Belief and Meaning for a detailed discussion of externalism and the ration-
ality problem. Bilgrami suggests that Burge has no satisfactory solution to this problem, and
that therefore externalism of Burge’s sort must be rejected.
It could be objected that the threat of conceptual fragmentation still remains. If every time individuals deviate from standard usage they should be reinterpreted and ascribed a non-standard concept, then no two people will ever share the same concepts. But this objection is not compelling. What leads to the concern about fragmentation is the claim that disagreements between the individuals and their communities typically indicate a conceptual difference. However, as I have argued above, this is not the case. Most disagreements do not show that there is a difference in concepts between individuals and their community. Only certain disagreements can be described as conceptual (and there are no sharp boundaries here), which is to say that only in certain cases do we have to conclude that the individual has a non-standard concept. Examples of such disagreements include cases where an individual rejects a trivial analyticity (e.g., holding that some bachelors are married) or where the disagreement is so radical and deep-running that it is better interpreted as a disagreement in meaning rather than in belief (e.g., not sharing any of our beliefs about arthritis except that it is a disease). In some cases, moreover, disagreement shows not that individuals have their own deviant concepts, but simply that they have mixed words up and are using a standard word in a non-standard way (an example, mentioned by Burge, is the man who thinks he has ‘orangutans’ for breakfast).

The dilemma that seemed to face the individualist is therefore a false one. There is no need to choose between giving up on reference on the one hand and accepting conceptual and referential fragmentation on the other. We must challenge the assumption that what has been described in the thought-experiment is a genuine case of ‘incomplete understanding’, a genuine conceptual error. Then the externalist conclusions are blocked: it can no longer be claimed that the word in question must express a different concept in the counterfactual community. And individualists can then safely grant that when we have a case of genuine conceptual deviation from the community, the speaker should be ascribed a deviant concept. Since this will happen only rarely, and not in the types of cases typically relied upon by social externalists, the threat of conceptual and referential fragmentation is averted and the individualist is free to move forward.15

University of Stockholm

15 Thanks to Henry Jackman, Michael McKinsey and Peter Pagin for valuable comments on earlier versions of the paper, as well as to the audience at the Twentieth World Congress of Philosophy in Boston (in particular, to Timothy Williamson), where a condensed version of this paper was presented. Special thanks to Kathrin Glüer for inspirational discussions on social externalism, and to an anonymous referee for this journal for helpful comments.

© The Editors of The Philosophical Quarterly, 2001