

NARROW CONTENT AND REPRESENTATION—or TWIN EARTH REVISITED*

Frank Jackson

Introduction

Intentional states represent. Belief represents how we take things to be; desire represents how we would like things to be; and so on. To represent is to make a division among possibilities; it is to divide the possibilities into those that are consistent with how things are being represented to be and those that are not. I will call the possibilities consistent with how some intentional state represents things to be, its content. There is no suggestion that this is the only legitimate notion of content, but for anyone who takes seriously the representational nature of intentional states, it must be one legitimate and central notion of content. To discover that DNA has a double helix structure is to make a selection from the various possible structures—or think of the rationale behind house-to-house searches by the police. I will focus on belief but our discussion is intended to apply *mutatis mutandis* to the other intentional states.

* I am much indebted to too many discussions to list but must acknowledge discussions (some supportive, some dissenting) with David Braddon-Mitchell, David Lewis, David Chalmers, Michael Smith, Jessica Brown, Daniel Stoljar, Martin Davies, Lloyd Humberstone and Michael Devitt. An earlier version of this material was given as the 2003 Patrick Romanell Lecture at the Central Division of the American Philosophical Association. I have kept some of the 'lecture' tone.

Our topic is whether Twin Earth teaches us that internally identical subjects' can have beliefs with very different representational contents merely by virtue of being in different environments. There is a piece of folk wisdom, exploited by the makers of The Truman Show and Total Recall, and which underpins brain in the vat science fiction stories, which says that one's environment matters for what one believes inasmuch as it has a *distinctive* effect on one. But the message of Twin Earth is alleged to be that differences in surroundings can make big differences in belief content without having distinctive effects on subjects. Many, a good majority I suspect, say that Twin Earth teaches us that belief content is broad, wide or anti-individualist, that how belief represents things to be is a partly environmental property of believers.

In this lecture, I argue that thinking of belief, and the words we use to express belief, in representationalist terms helps us see where the Twin Earth argument for broadness goes wrong. In particular, the representationalist way of thinking about content helps us understand three important matters: the first is centered content; the second is the sense of aboutness for which it is true that a difference in aboutness implies a difference in content; and, finally, it helps us understand the linguistic phenomenon of rigidity and its irrelevance to questions of belief content. Armed with these understandings, we can, I will argue, see where the Twin Earth argument goes wrong.

Here is the running sheet. I start with a preamble about the meta-philosophy of discussions of whether or not content is broad. I then enlarge a little on the

representational notion of content with an emphasis on the importance of centered content. We are then in a position to examine the Twin Earth case in its classic 'remote part of our world' form. I then discuss some objections. I finish with a discussion of belief and rigidity. This will involve us in the tricky question of how to recover what people believe from the sentences they use to express beliefs in cases where the sentences contain names, demonstratives and rigid designators more generally.

Meta-philosophical preamble

There is an important distinction between the first-order question of what the contents and references of certain of our words and our head states are, and the second-order question of what makes it the case that they are as they are. The debate between causal descriptivism and the causal theory of reference for proper names is in part the debate over whether the famous causal-historical link is part of the content or part of the *content-determiner* of a name. Casual descriptivism says the first; the causal theory of reference says the second.

The issue as to whether we can change contents merely by changing subjects' environments relates most directly to what makes it the case that our words and intentional states have the contents that they do—the second-order question. But any discussion of content that appeals to intuitions about possible cases, as so many do, relates most directly to what the contents and references of our words and thoughts are—the first-order question. For example, intuitions about what to say using the word 'water' about one or

another version of Twin Earth bear on the reference of the word 'water', not on how it got to have that reference.

The difficulty is that we are much more confident of our judgements concerning the references and contents of our words and beliefs than we are of our theories concerning why they are as they are. We all know what property the word 'square' refers to; we all know which person the word 'Gödel' refers to. This is despite the major disagreements about how those two words get to refer as they do. Although the second-order question is the prior question, we often do best to start with what our words and beliefs are in fact about and work back to what best explains this data. We don't want the more controversial to lead the less controversial.

Of course, there are some claims concerning the second-order question which are, or at least ought to be, non-controversial. The determination of content, both semantic and intentional, involves, in some broad sense, causal relations between subject and environment. But this is consistent with (representational) content being narrow in the sense of not being environmental, in the sense of not being a property we can change merely by transport from one environment to another.¹ Consider, for example, those who suffer from lactose intolerance. They have an unhappy causal relation to lactose. Nevertheless, lactose intolerance, and causal-dispositional properties in general, are narrow properties in the sense of being shared by

¹ Some writers tie the term 'representational content' to a co-variance answer to the second-order question. We—obviously—are not following this usage.

duplicates from the boundary in. What is striking about the Twin Earth case is that it appears to be a case where clear intuitions about answers to the first-order question tell us that we must move to a position on the second-order question that sees content as a property that comes and goes—perhaps with a time lag—with movements from one environment to another within a given possible world in the absence of changes under the skin.

One way to put the message of this lecture is that, when we look more closely at the Twin Earth case through the eyes of representationalism, we see that the intuitively appealing answers to the first-order question for the Twin Earth case do not support environmentalism about the second-order question. Twin Earth presented us with a research program: find an answer to the second-order question of how contents and references get to be as they are, what determines them, that explains why the contents and references so determined are in part a function of the environment in the way apparently revealed by Twin Earth. I am suggesting that we should think again about the case that launched the research program.

A note on terminology

I prefer to talk about whether belief content is an environmental property rather than whether it is broad or wide. This is to emphasise that the live issue, and the issue on the table here, is whether or not duplicates from the skin, doppelgangers,² in our world might differ in belief by virtue of a difference in

² I am taking the word to have become part of (philosophical) English,

their environment. In worlds where people think with major assistance from machines that they plug their brains into, doppelgangers will differ in what they believe (the skin will not be the pertinent boundary); and of course there is no plausibility at all in the idea that belief content is *inter-world* narrow—that is, that belief content tracks internal identity across worlds.³ The live issue is whether or not belief content is narrow in the sense in which, for example, causal-dispositional properties are narrow, and knowing that *P* and being selected to co-vary with *P*, for example, are not narrow.

Representational content and centered content

To believe that it will rain soon is to believe something about the arrangement of things and properties in our world; it is to believe that a certain portion of space-time will be full of rain. It is this way of thinking about the content of belief and intentional states generally that surfaces in talk of notional worlds⁴, possible worlds⁵, and positioned scenarios,⁶ and underlies the earlier point

³ For more what the live issue is, see Frank Jackson and Phillip Pettit, 'Some Content is Narrow', in Mental Causation, ed. John Heil and Alfred Mele, Oxford: Clarendon Press, 1993, pp. 259–82.

⁴ Daniel Dennett, 'Beyond Belief' in A. Woodfield, *Thought and Object*, Oxford: Clarendon Press, 1982.

⁵ As in, e. g., Robert Stalnaker, *Inquiry*, Cambridge, Mass.: MIT Press, 1984, and David Lewis, *On the Plurality of Worlds*, Oxford: Basil Blackwell, 1986.

⁶ Christopher Peacocke, *A Study of Concepts*, Cambridge, Mass.: M.I.T. Press, 1992. I mention positioned scenarios to remind that one of the best known advocates of accounts of content in terms of concepts uses the representationalist's notion for certain purposes.

that to represent is to make a division among possibilities: it is the rain-containing possibilities in the near future that get selected by the belief that it will rain soon. And it is this way of thinking about content that surfaces in the use of models, diagrams, computer graphics and maps to capture how one takes things to be.

I will use the familiar possible worlds way of framing the key issues but it will be important for what is to come that we be sensitive to the fact that very often what is represented is how things are *in relation to* that which is doing the representing. When I believe that there is an apple on my head, I believe something about how the region immediately above *my* head is. When t(win)Jackson, a duplicate from the skin (or brain) in of me, believes that there is an apple on his head, he believes something about how the region immediately above *his* head is. In consequence, my belief and his differ in truth conditions: one is true if and only if one of the regions is the relevant, apple-containing way, and the other is true if and only if the other region is the relevant way. Likewise, my thoughts and his will differ in reference. But this does not mean that our belief contents differ. We believe alike concerning how things are in the regions immediately above our respective heads—how things would have to be above our two heads for our beliefs to be true is the same for each of us. The same goes for the words we use to express our respective beliefs. How things would have to be above our heads for my token and his token of 'There is an apple above my head' to be true is one and the same.

The belief that there is an apple immediately above one's head is an example of an egocentric belief. Egocentric belief is ubiquitous. When we believe that it is raining, we typically believe that it is raining near where we *ourselves* are. Beliefs about the arrival of planes are typically beliefs about when the tarmac in front of *us* will contain a plane. Perceptual beliefs are typically beliefs about how we ourselves stand with respect to objects of various shapes, locations, colours and so on. This is why Christopher Peacocke's term 'positioned scenarios' is such a good one, provided we do not think of it as applying only to perceptual content.⁷

To capture egocentric content—of belief and of sentences—we need centered worlds. The content of my belief that I have a beard is the set of centered worlds that are the right way relative to the centers. The set of worlds that contain beards serves to capture the content of the belief that I am in a world with beards, but not that of my belief that I *myself* have a beard. Or consider John Perry's example: the belief that I myself am spilling sugar on the supermarket floor. Its content is a set of centered worlds. Each world has a sugar spiller as its center. And, as we learn from Perry and many others, this kind of belief is irreducibly egocentric.⁸ Divisions among possible worlds

⁷ See, e. g., Christopher Peacocke, *A Study of Concepts*, Cambridge, Mass.: MIT Press, 1992, ch. 3.

⁸ See, e. g., John Perry, 'The Problem of the Essential Indexical', *Nous*, 13 (1982): 3–21, Hector-Neri Castañeda, 'He*: A Study in the Logic of Self-Consciousness', *Ratio*, 8 (1966): 130–157, and David Lewis, 'Attitudes *De Dicto* and *De Se*,' *Philosophical Review*, 88 (1979): 513–543.

cannot capture the contents of these kinds of belief, and the contents of sentences in their role of expressing these contents; we need divisions among centered worlds. How an egocentric belief represents things to be is in general captured by a set of centered worlds, not a set of worlds *simpliciter*.

We can now say what we said earlier about the apple example in terms of centered worlds. The content of Jackson's and tJackson's beliefs that there is an apple above their respective heads is the same; it is the same set of centered worlds, the worlds with apple-above centers. But, as the center Jackson is actually at differs from the center tJackson is actually at, the conditions under which their beliefs are true differ. In the same way, the references of Jackson's and tJackson's beliefs may differ consistently with their having the same content in cases where we are dealing with centered content. This means that when differences in reference and truth are due solely to differences in centers, they are consistent with sameness of content. When we are dealing with centered content, truth and reference *simpliciter* are not truth and reference at the actual world; they are truth and reference at the actual world and the actual center, that is, at the world and center of the believer (or speaker).

Of course, centered worlds often appear in accounts of the way sentences like 'I have a beard now' can express different propositions depending on context of production. For example, 'I have a beard now' at context $\langle X, t \rangle$ is said to express the proposition that X has a beard at t . Equivalently, we have a function from pairs of centered worlds and worlds into truth values: in the

example, from a world with center $\langle X, t \rangle$ and world w into truth iff w is a world where X has a beard at t . It might be objected that this is the only role we need to give centered worlds. I am wrong to give centered worlds a central role in capturing content *per se*.

However, when I say that I have beard now, I am not saying, in the sense of representing, in the sense of saying how I take things to be, *who* I am and *when* I am speaking. Indeed, I may be in a dreadful state of confusion and have no idea who I am and when I am speaking but nevertheless have an egocentric belief, and a justified one, to the effect that I myself have a beard—whoever I am and whenever it is that I am having the thought that I am one of the bearded ones. It is this belief and its kin that require centered worlds in order to capture their representational contents. The set of worlds where X has a beard at t does not capture the content in question, no matter what values we give X and t .

In consequence, when we turn to the topic of the (representational) content of *sentences*, we need to acknowledge that sentences like 'I have a beard now' and 'There is an apple on my head here and now' have *two* contents. We can use sentences like these to give the egocentric or centered content of belief—a fact I have been exploiting—and this means that their contents, in one good sense, are given by sets of centered worlds: the very sets that we need for the beliefs whose contents they give. But it is also true that, for example, 'I have a beard now' at $\langle X, t \rangle$ expresses the proposition (in one good sense of 'proposition') that X has a beard at t , and the set of worlds where X

has a beard at t is not a set of centered worlds. It is the set of worlds where, as it might be, Jackson has a beard at noon, 14 February 2003.

In sum, my reply to the objection is that centered worlds have a dual role. One role is in modelling the dependence of the proposition expressed by certain sentences on context. The other role is that of capturing centered content, the content of egocentric belief, for example. Our interest is in the second role. The complication in the case of sentences is that a sentence like 'I have a beard now' serves to give the egocentric content of belief—witness what has already happened in this lecture—and so has a content given by centered worlds, but also expresses the proposition, as it might be, that Jackson has a beard at noon, 14 February 2003, a proposition that is true or false at a world regardless of center.

We are now in a position to look at Twin Earth in the remote place in our world version, the version Hilary Putnam famously gave us in 'The Meaning of "Meaning"'.⁹

Twin Earth examined

As we all know, on Earth the mostly but not invariably potable, odourless liquid kind that falls from the sky, fills the sea and so on, which is known to its English-speaking inhabitants under the name 'water' is H₂O. And, as all philosophers know, on Twin Earth—here thought of as somewhere in our

⁹ Hilary Putnam, 'The Meaning of "Meaning"' reprinted in *Mind, Language and Reality*, Cambridge: Cambridge University Press, 1975, pp. 215–271.

world but a long way away from us—the mostly potable, odourless liquid kind that falls from the sky and so on, which is known to its inhabitants under the name 'water' is XYZ. There is some H₂O on Twin Earth but it is black and tarry.

All philosophers also know roughly what this example has to do with the issue of whether content is broad. We are to suppose that the Earthians and Twin Earthians are identical from the skin in¹⁰ and are invited to agree that, nevertheless, their 'water' words and sentences differ in reference and truth conditions and, consequently, so do those intentional states, including belief, whose content we capture using the word 'water'. Our 'water' words, sentences and beliefs are about or refer to H₂O, and are true if and only if H₂O is the relevant way; whereas Twin Earthians' 'water' words, sentences and beliefs are about XYZ, and are true if and only if XYZ is the relevant way. When we say and believe that water is scarce, our saying and belief is about H₂O, and is true if and only if H₂O is scarce; when Twin Earthians say and believe that water is scarce, their saying and belief is about XYZ, and is true if and only if XYZ is scarce. But as Earthians are internally identical with Twin Earthians, the difference can only be explained by the difference in environments.

The trouble starts when we get less rough. The conclusion cannot be, and is not intended to be, that merely because XYZ is a long way away from us and

¹⁰ A complication is that H₂O is essential to life so that the idea that Earthians and Twin Earthians might be identical internally involves a big feat of imagination. But the points that follow could be made with less encumbered examples like tigers and schmigers, so we know no harm is done by ignoring

causally isolated from us, it is impossible for us to have beliefs about XYZ; ditto for Twin Earthians and H₂O. No one thinks it is impossible to have beliefs about stuff a long way away and causally isolated from one. Anyone who reads much astro-physics has many such beliefs. Of course, Putnam originally offered the example as showing us something about words. The extension to belief content came later in work by Tyler Burge that Putnam took on board.¹¹ But the point applies equally to words. Many words refer to stuff which is a long way away from us and causally isolated from us, and sentences made from those words are true if and only things are the right way a long way away and causally isolated from us.

The conclusion must, therefore, rest on some special feature of the Twin Earth example and, in particular, some special feature of our use of natural kind terms like 'water'. But when we review the obvious possibilities,¹² we see that no way of spelling matters out gives broad content theorists what they need. This means that we can be relaxed about which way of spelling matters out is the right way, and indeed it may well be that the matter is indeterminate or varies from person to person.

First, perhaps the word 'water' is a word for any watery substance, any substance that has most of the usual list of properties: clear, odourless, falls

the complication.

¹¹ Tyler Burge, 'Individualism and the Mental', in P. French *et al.*, eds, *Midwest Studies in Philosophy*, vol. 4, Minneapolis: University of Minnesota Press, 1979. There is, of course, a huge literature building on these papers.

from the sky, potable, and so on, and water beliefs are simply beliefs about watery stuff. In that case, 'water' in our mouths refers to XYZ every bit as much as to H₂O, and the beliefs we express using the word are about XYZ as much as they are about H₂O (but not the black tarry manifestation on Twin Earth). Likewise for Twin Earthians and H₂O. Not at all what broad theorists need, obviously.

Second, perhaps the word 'water' is a natural kind word in the sense that it refers to the unique natural kind that in some good number of manifestations, but not necessarily all, has the watery properties, and water beliefs are beliefs about the unique kind that in many manifestations has the watery properties. In that case, 'water' in our mouths fails to refer. There is no unique kind; there are two kinds that satisfy the specification: H₂O and XYZ. And our beliefs to the effect that there is water in the bath, and water falls from the sky when it rains, for example, are false. Similar remarks apply to Twin Earthians. Not at all what broad theorists need, obviously.

Third, perhaps the word 'water' is a natural kind word in the sense that it refers to any natural kind that in some good number of manifestations has the watery properties, and water beliefs are beliefs about any kind that in a good number of manifestations has the watery properties. In that case, 'water' in our mouths refers equally to H₂O and XYZ. Both kinds fit the bill. Similar remarks apply to Twin Earthians. Not at all what broad theorists need, obviously.

¹² Thanks to Asa Wikforss for convincing me of the need for this review.

Fourth, perhaps, the word 'water' is a natural kind word in the sense that it refers to the natural kind that in some good number of manifestations has the watery properties *and stands in such and such a relation to users of the word*, and water beliefs are beliefs about the kind that in many manifestations has the watery properties *and stands in such and such a relation to those who have the belief*. This seems to me the kind of view many supporters of Twin Earth have in mind. They often suggest that 'water' is a kind of demonstrative, not in the sense that it applies only to that which is demonstrated—they want 'water' to apply to water that no-one ever comes across—but in the sense of being a word which applies to whatever has such and such a relation to samples that are demonstrated by users of the word; and this makes 'water' a word for that which stands in the specified relation to the users of the word and havers of the belief.

We now have a spelling out that makes the references of Earthians' and Twin Earthians' utterances of 'water', and of the beliefs they give voice to using the word 'water', different. The problem with the earlier spellings out is that they do not pass this basic requirement on any version of Twin Earth that might sustain a broad content message. But our final spelling out ensures that the references are different. Our use of the word 'water' and our 'water' beliefs now do refer to H₂O alone. This is because H₂O, not XYZ, stands in the right relation to us as believers and word users; whereas Twin Earthians' use of the word 'water' and their 'water' beliefs refer to XYZ, not H₂O, because XYZ, not H₂O, stands in the right relation to them as believers and word users.

Moreover, our and Twin Earthians' beliefs and sentences will differ in the conditions under which they are true in the way supporters of drawing the broad content moral from Twin Earth claim. The beliefs we express with sentences like 'There is water near', 'The glass contains water', and 'Water is scarce', along with the sentences themselves, will be true, that is, true at the actual world, iff there is H₂O near, the glass contains H₂O, and H₂O is scarce; whereas the corresponding beliefs and sentences of the Twin Earthians will be true iff there is XYZ near, the glass contains XYZ, and XYZ is scarce.

However, we obtained these results by adding enough to the Twin Earth case to ensure that it involves centered or egocentric beliefs, and the words we use to express centered beliefs. The spelling out made it the case that our 'water' beliefs concern how certain stuff that stands in a certain relation to *us* is, and that Twin Earthians' concern how certain stuff that stands in a certain relation to *them* is. But, as we saw earlier, sameness of centered content goes along with differences in reference and truth in cases where we have different centers, and we and Twin Earthians are different centers with respect to 'water' thought and talk, on the spelling out of such thought and talk currently under discussion. The token beliefs and sentences in question are at different centers, and this means that the differences in respect to reference and truth do not show a difference in content.

In sum, the situation is as follows. Only if we ensure that the Twin Earth case is a case that involves centering, do we have a case where it is plausible that our, the Earthians', 'water' words and beliefs differ in reference and truth-

conditions from the Twin Earthians'.¹³ But if the case does involve centering, there is no longer support for there being a difference in content—in the representational sense that is our concern.

Replies to objections

When I have said things like this in the past, I have met many objections. Here is a partial list with my replies.

1. The first objection is that I have forgotten the distinction mentioned at the beginning between first-order and second-questions about content. Any treatment of the second-order question should allow that part of what determines contents and references are causal connections between heads and environments, and we Earthians are causally connected with H₂O, whereas Twin Earthians are connected with XYZ. This is the reason for holding that Earthians refer to H₂O, whereas Twin Earthians refer to XYZ; it has nothing to do with centering.

There are four problems with this answer. First, it would mean that we can no longer regard Twin Earth as a key part of the case for the causal theory of reference. The Twin Earth message would presuppose the causal theory rather than be an argument for it. Second, as far as causal facts go, it is as true that Earthians interact with and baptised mostly watery stuff, as it is that

¹³ Some insist that the Twin Earth moral does not turn on 'water' being somehow indexical. If (*if*) this amounts to insisting that it does not turn on 'water' language and belief being centered, this is bad news for drawing the

they interact with and baptised H₂O. The causal facts speak with a forked tongue. Thirdly and most importantly, we decide how to use our words, albeit that many of the decisions are implicit ones. As Kripke rightly says, we could have used 'Gödel' to mean the person who proved the theorem; it is an empirical fact about word usage that we don't.¹⁴ Equally, we could decide to use 'water' in a way that means it covers both H₂O and XYZ in the Twin Earth case. The causal facts don't stop us. But any plausible list of the decisions—and of the matters we left undecided—will have our word referring to H₂O and the Twin Earthians' word referring to XYZ, only if the decision is to use 'water' for that which stands in such and such a relation to users of the word in a way that would make beliefs expressed using the word 'water' centered ones. Finally, there are Twin Earth cases that cannot be handled by difference in causal connection. Suppose scientists on Earth come to believe in *elusium* on the basis of the need to fill a gap in *our* Earthian periodic table. They are right: elusium does exist but as its name suggests, it does not interact with us. Scientists on Twin Earth come to believe in what they call 'elusium'—'telusium' to us—on the basis of a gap in *their* periodic table. They are right as well. Elusium and telusium are, however, different elements. We can then develop a Twin Earth argument in the usual way, but the difference in truth conditions and reference between elusium and telusium sayings and beliefs cannot be laid at the door of the fact that scientists interact with elusium, whereas tscientists interact with telusium. Neither does either. The

moral, given the argument in the text.

¹⁴ Saul Kripke, *Naming and Necessity*, Oxford: Basil Blackwell, 1980. p. 91.

difference can only be laid at the door of centering. Scientists and tscientists each believe that there is an elusive element in their *own* world.

2. The second objection is that the difference in truth conditions which I grant under the supposition that 'water' utterances and beliefs have centered content is enough to make the broad content case: difference in truth conditions is sufficient for difference in content.

My reply is that it is not enough to make the broad content case *that is of interest*. Consider again the case where I and my twin believe that there is an apple above our respective heads. I said that we believe alike despite the difference in truth conditions. But suppose someone insists on saying that our beliefs are different. That difference would not be due to any difference in our surroundings; it would derive from the fact that we are different centers. The case would not support the thesis that content can be changed by changing environments while leaving subjects unaltered. A way to make the point vivid is to reflect on Tom who lives in possible world w_1 , and Dick who lives in w_2 . They both believe that some things are square. The conditions under which their beliefs are true differ: Tom's is true iff w_1 contains square things; Dick's iff w_2 does. But that does not betoken a difference in content.

3. The third objection is that what we learnt from Twin Earth is that reference in the case of natural kind words does not go by satisfaction of superficial properties—being watery—be they thought of as centered or not. It goes by standing in the same natural kind relation to some samples that have the superficial properties, and that is an outside-the-head fact.

My reply is that any plausible account of narrow content allows that reference can go by a multitude of factors including superficial properties and underlying nature as revealed by science.

The idea that narrow content theorists are restricted to superficial properties in the case of the word 'water' seems to come from the fact that when we acquire the word 'water', we typically know nothing of the underlying nature of the stuff we baptise; we are presented with some samples with the superficial properties and go from there. How then, it might be asked, can underlying nature rather than superficial properties be 'in our minds'? The answer is that presentation with samples with the superficial properties leaves open what we then go on to 'file away in our minds' concerning what to use the word 'water' for. We might use it for that which has the (superficial) properties, for that which has the underlying nature as revealed by best future theory of some of the stuff that has the properties, for that which is causally responsible for the properties in normal cases, for that which is similar to but does *not* have the properties, and so on and so forth.

4. The fourth objection is that there is a difference in *aboutness*, and that is enough to make the broad content case. Now, as we have noted, there are disambiguations of the word 'water', and correspondingly of which beliefs of the Earthians and Twin Earthians are under discussion, on which it is false that there is any difference in reference in their 'water' beliefs and sayings. On these disambiguations, there is no difference in aboutness. But let us suppose that we agree on the disambiguation that implies a difference in

aboutness, the disambiguation that involves centering. The objection might then be put as follows. "Forget about the fancy footwork with centering. The key point is that Earthian and Twin Earthian beliefs and words are about different kinds, and that means that they represent differently."

My reply is that although there is a sense of aboutness in which a difference in aboutness implies a difference in representational content, it is not plausible that it obtains in the Twin Earth case.

Suppose I hear a person through a partition wall talking on the phone. I can tell that it is a man speaking but have no idea who it is. Suppose, unknown to me, Fred is the person I can hear. Is my belief about Fred? It depends on how we read the relevant sense of aboutness.

There is an intensional sense of aboutness we can capture in terms of how things would have to be for my belief to be true. For my belief to be true, there must be one person talking on the phone in the next room. In this sense, my belief is about a single person talking on the phone in the next room. But, as I have no idea who that person is, it is not true that there is any particular person who has to be in the next room for my belief to be true. This is an intensional sense because the aboutness is captured by the fact that how things have to be for the belief in our example to be true requires that a person be on the phone, and this is consistent with there not in fact being anyone in the next room. In the intensional sense, my belief is not about any particular person, including Fred.

It is the intensional sense that is tied to the individuation of belief content (and linguistic content). A difference in aboutness in the intensional sense is a difference in how things have to be for the belief to be true, and such differences are what difference in content comes to from the representational perspective.

Are Earthian (centred) beliefs (and words) about H_2O , whereas Twin Earthians' are about XYZ, in the sense just noted as the one relevant to content individuation? The plausible answer is no. Consider people before Lavoisier and the rise of modern chemistry. If the Twin Earth argument works to show that our water beliefs are about H_2O in the sense that how things have to be for them to be true is that there exists H_2O , the same will be true for everyone with water beliefs, that is, just about everyone, before the rise of modern chemistry. But in fact the plausible view about people before Lavoisier is that they were in somewhat the same position as I am in the example of the person on the phone. They did not know what the watery stuff was or indeed anything about molecular structure, just as I do not know who is on the phone. In consequence, how they took things to be did not require that there be H_2O in order to be true. They only entered an epistemic state that required that there be H_2O in order for it to be true post the rise of modern chemistry. That is the sense in which modern chemistry enlarged our conception of what our world is like.

I know some will object that we Earthians have all along been in a belief state that represented that there is H_2O . What happened with the rise of modern

chemistry is that we came to know what we believed. Apart from being highly counterintuitive, I think this response misses a key point. Mere causal confrontation with K is not enough to represent that K exists in the sense of believing that there is K . For centuries we were confronted with electrons (as static and as lightning), with NH_3 in various forms, with proton donors (as acids), and so on, without being in belief states that represented that there are electrons, proton donors and so on. Confrontation in itself is nowhere near enough for belief in the representational sense. We need to attend lectures, carry out experiments, read books and all that, in order to acquire beliefs that represent that there is H_2O , NH_3 , proton donors and so on.

5. The fifth objection is that I have said nothing about the simplest way of putting the Twin Earth case for broadness. We Earthians report our belief that water is wet in these very words. But we cannot use these words to report the corresponding belief of Twin Earthians.¹⁵ The Twin Earthians can, but we have to use, for example, 'Twater is wet'.

My reply is that this is only true on the final possible usage of the word 'water' delineated earlier. On the first three usages, we can use our word 'water' to capture their beliefs. But when 'water' is used for that which stands in a certain relation to users of the word, there is a simple explanation of why Earthians cannot use their word for Twin Earthians' beliefs. To do so would be to

¹⁵ For a recent example, of this way of putting the Twin Earth case for broadness see Timothy Williamson, *Knowledge and its Limits*, Oxford: Oxford University Press, 2000, p. 53.

attribute to Twin Earthians beliefs about stuff that bears such and such a relation to us Earthians, whereas the beliefs they express with their word 'water' are about stuff that bears such and such a relation to them. As we might put it, to use our word 'water' for their beliefs would be to relate them to the "wrong" center. All the same, we believe alike in the sense that the set of centered worlds is the same for both of us. Analogy: you and I both believe alike when we believe that we ourselves have beards, but I cannot use 'I' for what it is that you believe.

Twin Earth and reference in other possible worlds

We have been discussing what Earthian and Twin Earth 'water' words and beliefs refer to in our world, and arguing that the plausible answers to these questions do not support there being a difference in belief content. But in many presentations of Twin Earth, the key claim concerns the reference of the word 'water' in other possible worlds. The claim is that 'water' in Earthians' mouths refers to H_2O in every possible world that has H_2O , regardless of whether or not H_2O has the famous watery features we listed earlier; similarly, 'water' in Twin Earthians' mouths refers to XYZ in every world that contains XYZ, regardless of whether or not it is ever watery in that world.

This is an *a posteriori* claim about word usage that requires, first, that 'water' is a centered word. The claim about the divergence in reference at other worlds of Earthian and Twin Earthian uses of 'water' is false on the first three possible uses of 'water' that we distinguished earlier (as is easy to see).

Second, it requires that 'water' be rigid in the mouths of Earthians and Twin

Earthians. But let's grant both claims. Let's suppose that we use the word 'water' for the *actual*, mostly watery kind that we are acquainted with. I am sure many of us use 'water' in this way. And let's make the corresponding supposition for Twin Earthians. It might seem that now we have a good argument for broadness cum environmentalism. For we now have a clear case where the reference of 'water' in our mouths at every world depends on a certain outside-the-head fact about us, namely, that H₂O is the mostly watery stuff we are acquainted with. Whereas in the mouths of Twin Earthians, for whom XYZ is the corresponding outside-the-head kind, the very same word, issuing from the very same kinds of heads, designates XYZ at every world.

I agree that we now have a good, indeed decisive, argument for broadness—for certain words. The 'other possible world' version of Twin Earth tells us that a certain kind of semantic externalism is true for many words, including 'water' as one among many examples (provided 'water' is used in the way under discussion). The reference of these words at non-actual worlds depends on facts about the actual world outside the heads of users of the words and, in consequence, so do the truth conditions of many sentences containing them. However, the reference of these words at *non-actual* worlds is a bad guide to the content of the beliefs we use them to express. I'll make the point with examples and conclude with some more general remarks.

Consider my belief about what is happening in the next room in the case given earlier where I have no idea who is on the phone. Suppose I decide to

give the person, whoever he is, that I can hear on the phone a name. This is clearly possible. We give things names all the time on the basis of our belief that they are the one and only thing satisfying some condition. (Of course, it is another question altogether whether, when we do so, the meaning of the name is 'that which satisfies the condition', or 'that which actually satisfies the condition', or the thing itself, or....) Suppose the name I give the person in the next room is 'Harry'. Doing this would not change how I take things to be on the other side of the partition. Belief is not that easy to change. But the truth conditions of sentences containing 'Harry' are very different from the truth conditions of sentences containing 'the person I can hear'. The precise nature of the difference is controversial but it is not controversial that there is a significant difference, and that part at least of the difference is that 'Harry' is a rigid designator, whereas 'the person I can hear' is not. A similar remark applies to demonstratives. Using the phrase 'that very person', perhaps accompanied by a special act of focussing attention on the source of the sound, would not change how I take things to be next door.¹⁶ Or consider what happens when astronomers name stars. They come to believe that there is an object which has properties like doing such and such to their instruments and their eyes, being in such and such a place, being a star, and so on. They give it a name. This act does not change how they take things to be, except of course that they now believe that there is one more act of naming than there

¹⁶ I take it that the role of demonstratives is typically to use the very fact that some item is being tracked by a mental representation as an element in securing reference but we will stick to issues raised by proper names here.

was a moment ago.¹⁷

The upshot is that we need to be very cautious moving from the semantics of sentences we use to report belief to what is believed in the representational sense in cases that involve rigid designators. The fact that the reference of rigid designators is broad does not imply that the beliefs we express using them have broad content.

Can we say something more constructive? This is a big question but here is a way of looking at how we might find belief content for beliefs expressed by the use of proper names. The attacks on the description theory of reference for proper names should not lead us to deny that users of names associate properties with those names in some good sense.¹⁸ Let me make the point with the Atlantis example. An expedition is formed to try and find the lost city of Atlantis. At a press conference the leader is asked how she will know when she has found Atlantis. She says that recent attacks on the description theory of reference have taught her that she need not, indeed maybe cannot, answer

¹⁷ For more on the principle that linguistic acts like naming and forming demonstratives do not change belief, see Gareth Evans, *The Varieties of Reference*, Oxford: Clarendon Press, 1982, p. 50, but I should emphasise that he would, I think, not agree with the use to which I'm putting the principle.

¹⁸ Properties, because the users may have no word for some of them, but in what follows we'll fudge the difference between properties and descriptions. Given our remarks about the ubiquity of egocentricity earlier, we should expect that many of the properties will be egocentric ones. I mention the point as some seem to think that the description theory of reference is tied to denying the irreducibility of the egocentric.

this question.

This would be a bizarre answer. She must have in mind what properties would allow her to conclude that she'd found Atlantis, and she must have in mind some putative properties of Atlantis when she decides, say, to sail South rather than North, and to start looking especially hard when the expedition gets near such and such a latitude. Essentially the same point can be made by reflecting on the great utility of proper names in allowing people to co-ordinate their behaviour. If people hear certain sentences containing 'Paris'—an example might be 'Paris is wonderful in the spring'—the probability of their ending up in Paris goes up substantially. But we find our way to Paris by putting together a series of co-ordinated intentional actions, guided at each stage by views about the distribution of properties. We do not have Paris-meters or Paris-thisness detectors in our heads. This is not to say that we consciously reflect on the properties guiding our intentional actions at any particular stage, though typically we can make a go of saying what they are if asked. The contrast is with the way sleepwalkers find their way around, and, maybe, with the way homing pigeons find their way home. When a pigeon finds home by using the alignment of the Earth's magnetic field, it does not know that it is using that alignment.¹⁹ However, when you go to the departure gate that has the word 'Paris' over it, you know which property you are using. Again, if historians of mathematics were ever to conclude that Gödel did not

¹⁹ I am indebted here to Simon Blackburn. There are a number of competing views about how pigeons find their way home, but I assume any plausible

prove Gödel's theorem, they would not be silent when asked how they identified Gödel, and they would not be silent when asked about the information about the distribution of properties that governed the waxing and eventual waning of their belief that Gödel proved Gödel's Theorem. (And if they were, we would pay no attention to their claim to have shown that Gödel did not prove Gödel's theorem.)

These reflections suggest a way of identifying the belief that lies behind a person's assertion of 'a is F '. The credence a person gives to the sentence 'a is F ' being true is always and only due to the credence they give to the thing which is such and such being F . This follows from the fact we lack "bare object detectors". This suggests that if we want to know how someone takes things to be in the sense of how they represent things to be when they produce the sentence 'a is F ', we should look to the belief of the form 'The so and so is F ' which best explains the credence they give 'a is F ' being true.²⁰

Of course many say that the content of the belief that a is F is an ordered pair consisting of a itself and the property of being F , or some such. This sounds

view gives the personal-level contents of their states a minor role.

²⁰ Direct reference theorists sometimes speak of the N role associated with proper name N . For recent example, see Robert Stalnaker, 'Conceptual Truth And Metaphysical Necessity', forthcoming. On some conceptions of this role, it may be that the 'so and so' above is the relevant role. Those who construe proper names as rigidified descriptions will likely hold that 'the *actual* so and so' is equivalent to the name in question, and that the recipe for finding the belief expressed reflects the fact that the belief expressed by 'The actual F is

very different from what I have just said, but it is possible to find common ground.²¹

We can find common ground if we say that this is another case where we need to distinguish two contents. One content of '*a* is *F*' is the ordered couple consisting of *a* and the property *F*. The other content is the set of worlds identified earlier in terms of the source of our credence in the sentence's truth. It is the set of worlds where 'The so and so is *F*' is true, for the value of 'The so and so is *F*' that gives the source of our credence. It is, if you like, the representational content of the sentence, and is the content of the belief about how things are that we use the sentence '*a* is *F*' to make public.

Another example is the sentence 'There is water'. The belief about how things are that English speakers use this sentence to make public is very different from the belief they use the sentence 'There is H₂O' to make public, despite the fact that, in *one* sense of content, they have the same content because they make the same division among possibilities in *one* sense (assuming rigidity). To say otherwise is to make a nonsense of the epistemic significance of the change in the sentences we produced before and after Lavoisier. But it is plausible that our confidence in the truth of 'There is water' marches in step with that of 'There is stuff which is thus and so'. Perhaps, as David Chalmers and I have argued elsewhere, the 'thus and so' is something like 'the generally watery stuff of our acquaintance' (see references in fn. 22). If

G' and the '*The F* is *G*' is one and the same.

²¹ I *think* I am agreeing with at least some of what Evans says, *loc. cit.*

this is right, the belief about how things are that we make public with the sentence 'There is water' is that we inhabit a world where there is watery stuff, and the relevant content is the set of worlds where there is watery stuff. If you think, as I do, that most users of the word 'water' include an egocentric element, the content will, more strictly, be a set of centered worlds.

If you go further—as I would but many would not—and conclude that we should think of the name '*a*' as a rigidified description, that is, as equivalent to 'the actual so and so', we can put the two-content doctrine in terms of the distinction between *A*-intensions and *C*-intensions (or primary and secondary intensions in Chalmers's terminology). The *C*-intension of '*a* is *F*' is the set of worlds where the object *a* is *F*. The *A*-intension of '*a* is *F*' is the set of worlds *w* (or centered worlds) such that 'The actual so and so is *F*' is true at *w* (or at $\langle c, w \rangle$) under the supposition that *w* (or $\langle c, w \rangle$) is actual; that is, the set of worlds (or centered worlds) where 'The so and so is *F*' is true. It is the *A*-intension and not the *C*-intension that gives the content of '*a* is *F*' in the sense of the content of the belief about how things are which we give voice to using the sentence '*a* is *F*'.²²

²² The terminology of *A*- and *C*-intensions is the one I favour in *From Metaphysics to Ethics*, Oxford: Oxford University Press, 1998. Chalmers mostly favours the terminology of primary and secondary intensions, see David J. Chalmers, 'Précis of *The Conscious Mind*', *Philosophy and Phenomenological Research*, LIX, 2, 1999: 435–438, but see also 'The Components of Content', in Chalmers, ed., *Philosophy of Mind: Classical and contemporary readings*, New York: Oxford University Press, pp. 608–633. An important difference in

Some who are happy to talk of truth at worlds worry about the talk of truth at worlds under the supposition of actuality (and *mutatis mutandis* for truth at centered worlds under the supposition of actuality) in accounts of A-intensions. Truth of sentences at worlds in the relevant sense is truth at worlds with the sentences having the meaning they actually have. But, they worry, doesn't that require us *not* to vary which world is actual (*mutatis mutandis* for centered worlds)? However, what it is to be true at a world under the supposition that it is actual is cashed out in terms of what it is to be true at a world *simpliciter*. For example, 'The actual *F* is *G*' is true at *w* under the supposition that *w* is actual is true iff 'The *F* is *G*' is true at *w*, and the key feature of this example, namely, that what follows the 'iff' uses truth at *w*, not truth at *w* under the supposition that *w* is actual, applies across the board (*mutatis mutandis* for cases that involve centered worlds).

I think we should not be surprised that some fancy footwork is required to capture how a subject believes things to be when they assert sentences containing rigid designators. How I take things to be is none other than how I take things to be actually: representational content is all the ways things might

our views is that I see the distinction as one for language and not for thought—for more on this, see Jackson, 'Responses' in the panel discussion of *From Metaphysics to Ethics, Philosophy and Phenomenological Research*, LXII, 3, 2001: 653–664, pp. 662f—whereas Chalmers gives the distinction wider application. Often a sentence's A-intension is a set of centered worlds: the set of centered worlds such that the sentence is true at the centered world under the supposition that center and world are actual, but this complication is ignored in the text.

actually be consistent with how I take things to be—that is, the set of worlds where a sentence's truth is consistent with that world being the actual world (*mutatis mutandis* for cases that need centered worlds). And that precisely is the *A*-intension of the sentence. With sentences that contain no rigid designators, the *A*-intension is one and the same as the *C*-intension, but not in the general case, where we need the fancy footwork.

The Australian National University