

REFERENCE AND DESCRIPTION FROM THE DESCRIPTIVISTS' CORNER

FRANK JACKSON

The Australian National University

I

Disagreement is endemic in philosophy. Even so, the extent of disagreement over reference and description is surprising. It is surprising because there is very substantial agreement over the data. We mostly agree that names are rigid designators, unlike typical definite descriptions; that 'Gödel' is not short for the person who proved the incompleteness of arithmetic; that water on Twin Earth is H₂O, not XYZ; and that causal history and naming ceremonies are very often an important determinant of the reference of proper names and names more generally. Indeed sometimes the very same data—intuitions about what refers to what—are taken by opponents of the description theory as making a powerful case against the description theory but by its supporters as serving to elucidate the descriptions that determine reference.



I think the extent of the disagreement is in part fuelled by important differences in the way the key issue on the table is approached. In this discussion I will sketch (all that's possible in the space available) the way description theorists (or this one anyway) approach the key issue on the table and how this shapes the ensuing debate. This will allow me to identify some of the points of difference with Scott Soames's sustained criticism in *Reference and Description* of the description theory of reference and of two-dimensionalism.¹

II

Description theorists start from the fact that language is a system of representation, or much of it is anyway. This is something Soames and I agree about, while seeing the implications of the point very differently, as we will see.

The representational nature of language is obvious from the role of sentences we understand in conveying putative information about how things are. Sentences in English (to stick with the language in common to readers of this journal) are an enormously valuable source of information about how our world is: think of how much about our world we learn from reading and listening, and that is only explicable if very many of the sentences in English represent that things are thus and so in ways that are accessible to us. This is as true for sentences containing names as it is for sentences containing adjectives. The fact that 'The conference is in London' represents that things

1. *Reference and Description: The Case against Two-Dimensionalism* (Princeton University Press, 2005).

	P	H	I	B	4	3	1	Operator: LinPing		Dispatch: 10.11.06	PE: Elizabeth Howdill
	Journal Name				Manuscript No.			Proofreader: Li aijiao		No. of Pages: 10	Copy-editor:

are a certain way is essential to the way hearing or reading the sentence enables those who want to go to the conference to end up in the right place. There is no difference on this score from the way that reading or hearing the sentence 'The treasure is in the *square* box' assists in finding the treasure. Moreover, the way a sentence represents things to be is a contingent, a posteriori matter. We all know how 'There are circles' represents things to be. It represents that there are circles, things with the same shape as 'O'. But it might have represented that there are squares. Furthermore, our decisions to use words and sentences in certain ways play a key role in making it the case that words and sentences represent as they do. No-one forced us to mean what we mean.

Descriptivists draw three morals from all of this. The first is that a theory of reference for names is an account of the contribution that names make to how sentences containing them represent things to be. What contribution, to illustrate, do the names 'water' and 'Sammys' make to how we represent our world to be when we utter 'There is water nearby' or 'Sammys is the best cheap restaurant'? There is of course a degree of prescription here. 'Reference' is a term of art in current philosophical discussions and there are many topics that might be called the theory of reference for names. Also, we should not assume the contribution 'Sammys' makes in a fairly simple sentence of the form 'Sammys is . . .' is the same as that it makes in, for example, belief reports. The second moral is that how we represent things to be had better be accessible to us. More on this anon. The third moral is that our account of the reference of names had better take into account the point about choice. We will start with this last point.

III

Saul Kripke makes the point about choice in *Naming and Necessity*,² p. 91:

The picture that leads to the cluster-of-descriptions theory is something like this: . . . one determines the reference for himself by saying—'By "Gödel" I shall mean the man, whoever he is, who proved the incompleteness of arithmetic'. Now you can do this if you want to. There's nothing really preventing it. You can just stick to that determination. If that's what you do, then if Schmidt discovered the incompleteness of arithmetic you *do* refer to him when you say 'Gödel did such and such'. (p. 91)

But now we face two questions. First, if we could have decided to use 'Gödel' for whoever proved the incompleteness of arithmetic but didn't, what decision did we in fact make concerning the use of 'Gödel' that meant that it did not mean whoever proved the incompleteness of arithmetic? The second question concerns our evidence that 'Gödel' does not mean whoever proved the incompleteness of arithmetic.

The second question is perhaps rather embarrassing for philosophers. It is a question in empirical linguistics and few philosophers have done much

2. Blackwell, 1980.

empirical research on the ways the typical English speaker uses ‘Gödel’, and a similar point applies to ‘water’, ‘Aristotle’ and all the other examples that figure prominently in the debate over the reference of names. We might, for example, have used ‘water’ in a way that meant it would have applied to XYZ in the Twin Earth scenario. What is our evidence that we don’t?

I think there is only one way for philosophers to meet the charge that they have treated something that is clearly an empirical question as if it were an *a priori* matter. It is by arguing that in fact a good deal of fieldwork has been carried out, though not packaged as such. The very describing of possible cases in seminars and publications—the describing of the imaginary case where Schmidt was cheated of the glory by Gödel, the describing of Twin Earth and of smigers, and so on—followed by a good degree of endogenous agreement in the relevant intuitions—that Schmidt is not Gödel, that XYZ is not water, that smigers are not tigers, and so on—is empirical evidence. The degree to which the agreement is endogenous might be questioned in some cases. David Lewis may well be right that ordinary users of English are undecided about whether or not ‘water’ refers to XYZ on Twin Earth, and that the consensus among philosophers that it doesn’t is tribute to Hilary Putnam’s advocacy.³ Be this as it may, it is an empirical question as to whether or not ‘Gödel’ is used to refer to the theorem prover and ‘water’ is used in a way that means it does not refer to XYZ in the Twin Earth case, and the famous intuitions enter the picture as our best reply to the charge of treating an *a posteriori* issue as if it were an *a priori* one.

IV

But now we have a major issue for opponents of the description theory of reference. If they cannot dismiss the intuitions, how do they meet the charge that, as the intuitions concern what refers to what in described cases, the intuitions serve to elucidate the descriptions that determine the reference of names and so cannot be part of a refutation of the description theory? Soames indicates strong disagreement with this line of thought in a number of places in the book. Here’s a passage from p. 36:

For example, Frank Jackson argues that
Our ability to answer questions about what various words refer to in various possible worlds . . . is common ground with critics of the description theory. The critics’ writings are full of descriptions (*descriptions*) of possible worlds and claims about what refers, or fails to refer to what in these possible worlds. . . . But if speakers can say what refers to what when various possible worlds are described to them, description theorists can identify the property associated in their minds with, for example, the word ‘water’: it is

3. David Lewis, ‘Reduction of Mind’, in S. Guttenplan (ed.), *A Companion to the Philosophy of Mind* (Blackwell, 1996), pp. 412–431. We will, however, presuppose the usual view among philosophers in what follows.

the disjunction of the properties that guide the speakers in each particular possible world when they say which stuff, if any, in each possible world counts as water. This disjunction is in their minds in the sense that they can deliver the answer for each possible world when it is described in sufficient detail, but is implicit in the sense that the pattern that brings the various disjuncts together as part of the, possibly highly complex, disjunction may be one they cannot state.

This is a remarkable defence. If correct, it might seem to suggest that description theories of reference are virtually guaranteed, a priori, to be irrefutable, since any refutation would require a clear, uncontroversial sketch of a possible scenario in which *n* referred to something *o* not satisfying the description putatively associated with *n* by ordinary speakers like us (or failed to refer to the thing that was denoted by this description) whereas the very judgement that *n* does not refer to *o* in this scenario (or does not refer to what the description denotes there) would be taken by Jackson to demonstrate the existence of a different, implicit description in our minds that successfully determines reference, whether or not we can articulate it.

The trouble for opponents of description theories is that they must allow the relevance of the intuitions, and given that, I do not see what Soames's reply is. Yes, the conclusion is a strong one, but noting that is not a reply to the argument.

V

The first question raised by the quotation from Kripke—what decision did we make that led to 'Gödel' not meaning whoever proved the incompleteness of arithmetic—raises some difficult issues about explicit versus implicit agreements. Some philosophers enter into explicit agreements to use 'Julius' to mean whoever actually invented the zip.⁴ But words for which there is some explicit agreement are very much the exception. What then is the *implicit* convention governing our use of names in English?

This is a big topic quite outside the word limit of a discussion but two contentions are plausible. The first is that because it is something *we* agreed to, even if implicitly, it is something available to us. The second is that a key part of what we agreed to in many cases is that, as part of the project of transmitting information about particular objects, names get attached to the objects in question and are then used in ways that make the sentences containing them part of information-preserving causal chains. The useful information provided by a guidebook on Canberra that includes the sentence 'Sammys is the best cheap restaurant' is that those wanting a good cheap meal should go to the restaurant named 'Sammys'. Someone who doesn't know that that is the sort of information provided by the appearance in

4. Following, of course, Gareth Evans, *The Varieties of Reference* (Oxford University Press, 1982), Ch. 1.

guidebooks of a sentence like that has a defective grasp of the way names work in English in much the way that someone who reads on a conference program 'The next session is in room 4277' but does not know to look for a room with '4277' in some way attached to it does not understand how room numbers work. It is common knowledge that sentences of the form 'Sammys is . . . ' provide useful information, first, because something was in the past given the name 'Sammys', and, second, because the way names typically get used means that the sentence token in the book lies at the end of an information-preserving causal chain.⁵ Of course there are exceptions. 'Julius' as used by many philosophers is an exception, as are names for things outside the light cone.

This is a second point of disagreement with Soames. Responding to an argument of mine turning on the point that words do not get their reference by magic and therefore that there must be some story to be told about what makes it the case that a name refers as it does, Soames distinguishes (pp. 182–3) "(i) the facts that originally brought it about that *water* stands for what it does, and that have sustained the reference of the word since it was initially established, and (ii) the facts about the meaning of *water* that speakers must master in order to understand the word". And he makes a similar distinction on p. 185, this time in the context of some views of Kripke's on proper names:

... 'reference fixing by description' and 'reference fixing by historical chain' are not on a par. Reference fixing by description is, as Kripke standardly understands it, a matter of linguistic rules that competent speakers must master if they are to understand the relevant terms. By contrast, when one speaks of reference fixing by historical chain, one is alluding to certain foundational facts of the sort (i) above that bring it about that a name or natural kind term comes to refer to what it does, and that play important roles in determining that the original reference is preserved over time. These foundational facts are not parts of the meanings of names or kind terms of which competent speakers must be aware. Speakers have to know what the words of their language mean and refer to, and they must know how to use them. They don't have to know how the conditions for their proper use arose and are sustained.

As against this, I insist that someone who doesn't know that typically names get given to things and does not understand the causal informational role of names does not understand their role in English. The point of sentences like our guidebook sentence rests on the folk knowing facts like these. Indeed, it is precisely this fact which explains why, when asked what a name refers to in various scenarios, they are so influenced by the causal historical facts stated in the scenarios. Or take the sentence mentioned near the beginning: 'The conference is in London'. The way this sentence enables those who hear it or read it, and who wish to attend the conference, get to London relies on two

5. For more in this vein see Frank Jackson, 'What are Proper Names For?', in Johann C. Marek and Maria E. Reicher (eds.), *Experience and Analysis* (hpt-öbv Vienna, 2005), pp. 257–269, and John Searle, *Intentionality*, (Cambridge University Press, 1983), Ch. 9.

facts: the knowledge of the hearers and readers that (i) they should go to somewhere called 'London', and (ii) that various other sentences they come across like 'London is served by the following airlines . . .' are a reliable source of information about London, something resting, as English speakers know, on information-preserving causal chains running from London to the sentence tokens they come across. The way the folk use sentences containing proper names in finding their way around our world, and as a source of knowledge about our world, shows that they know about naming practices (baptisms) and about how our response to them generates information-preserving chains. This is how they know that the appearance of 'Real Madrid won last night' in the morning paper gives them information about a happening spatially and temporally distant from the token. Of course I'm not denying that one might refer to London or Real Madrid successfully without knowing the facts about information chains and naming. As a description theorist I allow that any uniquely satisfied description will do to secure reference (as in effect does Kripke in the quotation near the beginning), but one's understanding of the role of such names would not then be the understanding the folk give those names.

VI

I now turn to the implications of the second moral mentioned near the beginning—that how we represent things to be had better be accessible to us. Here is where two-dimensionalism has to come into the story, or so it seems to me.

The sentence 'There are four chairs and one table in the next room' provides putative information about how things are in that room by virtue of our *knowledge* of which of the various ways things might be in that room are consistent with the sentence. It is not enough that the sentence represents in the sense of effecting a division of the possible ways. In addition we need to know what that division is. Sentences in languages I do not understand effect divisions in the possibilities but are informationally useless to me because I do not know what the division is. There are two parts to this picture: to represent is to make a division in possibilities, and the division needs to be known if that which represents is to provide information. If you grant this plausible picture, you more or less have to embrace some form or other of two-dimensionalism, as we'll now see.

If you say 'There is something round nearby', I know how you are representing things to be and I could arrange matters so that things were as you were representing them to be: I'd find something round and place it near to you. In doing this, I would make the sentence you produced true (unless it was already true of course). So, how you represent things to be and the possibilities at which your sentence is true are one and the same. Unfortunately, for a whole range of sentences this answer is badly astray.

Here is an example. You believe that you were robbed by the same (unknown) person on both Monday and Tuesday nights while you were out. You might use 'The person who robbed me on Monday night robbed me on Tuesday night' to capture how you take things to be, or you might use 'The very (in the sense of the actual) person who robbed me on Monday night

robbed me on 'Tuesday night'. No difference in how you represent things to be: in both cases you are affirming that it is the same person both times without saying, or indeed knowing, who it is. But the worlds at which the sentences are true are different. The first sentence is true at worlds where it is the same person both nights. The second sentence is true at worlds where it is the same person both nights *and* that person is the one responsible in the actual world. The first set of worlds gets the content of what you are affirming to be the case correct; the second set wrongly makes the content out to be much richer than it in fact is.

This is one example among many where the set of possibilities where a sentence is true is not how the user of the sentence is representing things to be, where, as we'll say it, the set of possibilities where the sentence is true does not give the representational content of the sentence. We need to look for some other set of possibilities—a second dimension—to capture representational content. Another case is the sentence 'I have a beard'. Produced by x at t , it is true at worlds where x has a beard at t , but x is not saying anything about who he is or when it is. What is more, there are well known reasons for holding that sentences like 'I have a beard' represent not how a world is, but how a part of a world is, and so call for sets of centred worlds to capture their representational content. Another example is the sentence 'There is water nearby', especially when produced before the rise of modern chemistry. The belief about how things are voiced by the production in 1700 of 'There is water nearby' is not that there is H_2O nearby despite the fact that the worlds at which the sentence is true are those where H_2O is nearby. (Some dissent here, saying that externalism teaches us that they believe that there is H_2O nearby, albeit that they do not know this, and therefore that what they affirm when they produce 'There is water nearby' is that there is H_2O nearby. But if that's right, the philosophers of 1700 should have been telling the folk that they did not know what they were saying, and what belief they were expressing with their 'water' utterances, but fortunately Lavoisier would soon help them out. Surely, this is a last resort position.) Finally, we can make the same point with an example Soames thinks of as making serious trouble for me. He urges that it is very implausible to hold that he, Soames, does not know the proposition expressed by sentences containing his name, 'Scott Soames', and those of his two sons, 'Greg Soames' and 'Brian Soames'. But what Soames plausibly does know is what he is representing about how things are when he uses such sentences. But the view of mine he is attacking is that he does not know the proposition expressed *in the sense of the worlds at which the various sentences are true*. And that view of mine *is* plausible: fathers know their children and themselves well, but do not know their essences, and that's what's required to know the worlds at which, say, to borrow from his example on p. 168, 'The Soames's are all composed in part of H_2O molecules' is true. The example, far from being a problem one for me, makes the point nicely that we must distinguish how things are being represented to be, a sentence's representational content, from the worlds at which the sentence is true in cases like these.

The challenge, accordingly, is to find the right set of possibilities for representational content, and there are two different ways one might go here. The

first is to argue that we should distinguish metaphysical possibilities from epistemic possibilities, and identify the representational content with the set of epistemic possibilities (possibilities, not possible worlds in case the sentence is one that calls for centred worlds) at which the sentence is true. The idea is that we avoid the result that, for example, ‘There is water nearby’ represents that there is H₂O nearby, because it is epistemically possible that water is not H₂O and so the set of epistemic possibilities at which water is nearby is not one and the same as the set of epistemic possibilities at which H₂O is nearby. The other way to go is to work with a generic or vanilla notion of possibility, call it metaphysical possibility, and look for a different set of metaphysical possibilities from those at which the sentence is true to be the representational content of the sentences in question. Instead of carving from a different set of possibilities, we get the second dimension by carving differently from the same set.

I have argued in a number of places for this second option, suggesting that the needed set is the set of metaphysical possibilities whose actuality is consistent with the sentence’s truth, the set at which the sentence is true under the supposition that the possibility is actual. This is the primary intension or *A*-intension on some spellings out of that notion (there are others). I won’t rehearse my arguments here but an example Soames discusses on pp. 198–199 enables me to sketch one reason I favour the second option. He takes the case of a paperweight that is in fact made of wood but, for all you know, might be made of metal or plastic. He describes this as a case where it is metaphysically impossible but epistemically possible that the paperweight be made of plastic, saying “In my view, the natural, default position is that, of course, there are metaphysically impossible but epistemically possible world-states—ways the world could not genuinely be which we cannot know apriori that it isn’t.” But what, precisely, is the *world*-state in question that is both metaphysically impossible and epistemically possible? Let’s ask, first, what is supposed to be metaphysically impossible. Presumably, the possibility of that very paperweight being made of plastic. But, as being made of wood is an essential property of that paperweight, that world-state is that of something made of wood being made of plastic, and that’s epistemically impossible as well as metaphysically impossible. Let’s ask, second, what is supposed to be epistemically possible. Presumably, the possibility of that very paperweight being made of plastic, and as being made of plastic is not an epistemically essential property of that paperweight, that world-state is that of a paperweight very like the one in question without being made of plastic. But that’s metaphysically possible as well as epistemically possible. When you dig into the case, we do not have a single world-state that is metaphysically impossible and epistemically possible but two world-states, one that is both metaphysically and epistemically impossible, and another that is both metaphysically and epistemically possible.

VII

We can now place the debate over two-dimensionalism in a rather different light from that it receives in Soames’ book. He paints a picture in which the

big driver for its philosophical application to the debate over reference and description is a liking for the view that that which is metaphysically necessary is also a priori. But if what we've just been saying is right, everyone had better allow that we need two dimensions on pain of getting the representational content of many sentences badly wrong. It is then a matter for further discussion how we characterize the two dimensions.

What is more, we need a principled way to get the representational content of, for example, 'There is water nearby' right. My own suggestion is to think of 'water' along the lines of 'the actual watery stuff', to think of it as a rigidified description. What is important is not the detail—'watery stuff' was always a rough schema that might be filled out in many ways to allow for variability between users of the word and a good degree of indeterminacy—but the way the suggestion allows us to find, in the needed principled way, the additional set of possibilities. These are, on the approach I like, the possibilities in which there is watery stuff nearby, for these are the possibilities whose actuality is consistent with the truth of 'There is water nearby', on the rigidified description account of 'water'. Although Soames is critical of this suggestion, his own words suggest something similar. Here is what he says about what is needed to understand the word 'water' in English:

... an otherwise competent English speaker is counted as a competent user of the word, if he or she knows that *water* is a term that stands for some natural kind that determines its extension at different world-states—even if one doesn't have any reliable way of describing that kind, other than the kind that the word stands for in English. Perhaps one also has to have some idea of what type of kind it is—i. e. that it has something to do with physically constitutive characteristics—and that the stuff in question sometimes comes in liquid form. (pp. 183–184)

Suppose what he says here is right, under what circumstances would it be correct for someone to use the word 'water' to describe how they take their world to be in a sentence like 'There is water nearby'? Surely just when they believe that there is "some natural kind that determines its extension at different world states ... [is] the kind the word ['water'] stands for in English ... has something to do with physically constitutive characteristics ... and ... sometimes comes in liquid form" is nearby. This is exactly the kind of account I like. Soames has given us his version of 'watery'. One might or might not like the detail of his account but the principle is the same.

VIII

Two final matters. The first is a point of clarification about my views. Soames says (p. 172) that "The main feature of strong two-dimensionalism about which he [Jackson] is less than fully explicit involves the semantics of attitude ascriptions. ... For the strong two-dimensionalist these (standardly) report relations between agents and the primary intensions of their complement

clauses”. He says that I seem “tacitly to assume this” on the basis that I say that the primary (*A*-) intension of a sentence affirmed by someone often best captures what they believe. For the record I should say that I don’t assume the view he calls standard for strong two-dimensionalists. I think the semantics of attitude ascriptions is a highly complex matter due to, *inter alia*: the tendency for reporters to incorporate their knowledge into the report—as in ‘Poor Fred, he believes that the person who is in fact undermining his chances for election is his biggest supporter’, the tendency for names in belief reports to take wide scope, and potential unclarity about whether it is the intension of the reporter or of the believer that governs the choice of the sentence to be the complement clause.

A more substantive point is the role of centred content. I earlier mentioned this as important for seeing why the worlds at which a sentence is true often don’t give its representational content; we need centred worlds instead. It is also important for seeing why an objection Soames regards as an especially powerful objection to the description theory of reference fails. Here’s the objection, quoting from p. 25:

Another example [with anti-descriptivist implications] involves Kaplan’s identical twins, Castor and Pollux, raised in qualitatively identical environments to be molecule for molecule identical and so, presumably, to associate the same purely qualitative descriptions with the same terms. Despite this, each refers to himself, and not the other when he uses *I*.

I suspect that there are two objections being run at once here. One is that difference in reference with sameness of associated descriptions refutes the description theory, and does so without any fussing about which descriptions determine reference on any given version of description theory. However, the example is very obviously one involving centred content. The reference of ‘I’ is a function of world *and* centre, just as ‘here’ typically is. And difference of reference is compatible with sameness of associated descriptions when centres are different. The second objection is that the associated description cannot be given in purely qualitative terms. But what is true is that we cannot capture the content of a saying like ‘I am happy’ in purely qualitative terms. This is the reason we need centred content to capture what Castor, say, is representing about how things are in such sentences. If a purely qualitative translation were possible, we could capture the content in terms of a division among complete ways things might be, possible worlds, and there are powerful arguments showing that this cannot be done. Nevertheless, the relation between each token of ‘I’ and that to which it refers is purely qualitative: it is being caused in a certain way. When you hear someone at a seminar utter ‘I would like to ask about your second objection’, you *qua* member of the folk, know what the reference of that token of ‘I’ is, for you are (typically) able to observe the relevant facts about the causal origin of the token of ‘I’.⁶

6. I am indebted to many discussions over many years, especially with David Lewis, Michael Smith, David Braddon-Mitchell and David Chalmers.