Recognitional Capacities and the Determinacy of Meaning

Robert West

Introduction

The problem which I want to address in this paper is whether one has to be committed to absolute determinacy of reference in giving a satisfactory account of the meaning of natural kind terms. Kripke and Putnam argue that the only way to get the right truth-conditions for our use of natural kind terms is to include elements of the world as part of the meanings of such terms, so that meanings aren't in the head.

I want to suggest that there are at least two distinct notions of "meaning" and that the most important one, that is to say the notion most commonly deployed in ordinary discourse, really is in the head, but also that it is not the sort of thing which can produce determinate truth conditions on every occasion of use. I will argue that there is nothing wrong with the theories of Kripke and Putnam - or at least with the general approach of such theories - but that we should be clear about what the status of such theories is exactly. The postulation of a division of linguistic labour is correct for communities which have the relevant body of experts, but their pronouncements should not be seen as granting an absolutely determinate extension for natural kind terms. It is wrong to think that science allows us to engage with fundamental divisions in nature in a way that grants determinate extension to natural kind terms in a way that is not theory-relative.

I will claim that we can satisfy our desire for philosophical explanation by changing our perspective. A satisfactory story can be told about the use of some referring terms and natural kind terms which reveals that it is a mistake to demand determinacy of reference on all occasions. Thus in some instances, which will be enumerated below, there is no matter of fact about what *S* means by "water". In such cases what we should do is describe what the individual does with that term, including what they would say and do in counterfactual situations. Such an enterprise is envisaged to be empirical, and analogous to taking a complex machine and either seeing how it reacts to actual stimuli or gaining an understanding of how it functions such that one can then specify how it *would* react to a stimulus *were* it to be presented. Once this is accomplished no more can be done to fix meaning and reference. Now in some cases - perhaps the majority - such a description will yield determinate meaning and reference, i.e. a set of truth-conditions. However, it is my hope that by spelling out how the empirical explanation might go it can be made clear how it can be tenable to acknowledge a kind of meaning which is not fully determinate. In essence, what I want to deny is that for natural kind terms to be usable they must latch onto a division which exists independently of us in nature.

This approach owes something to Wittgenstein, for several reasons. Firstly, it was his insight that meaning does not have to be exact ("everywhere bounded by rules"[1]) in order to function satisfactorily. Secondly, Wittgenstein propounded the idea that an explanation of meaning should

---

[1] Wittgenstein, L. (1953) §84.

proceed via an investigation of use. However, I think it is unsatisfactory to leave things at this point, and I want to adopt a richer account of linguistic behaviour, and most importantly of its underlying cognitive mechanisms, using the notion of a recognitional capacity. With this in place it might be possible to give an account of language use which is not in itself philosophical, but which nevertheless allows us to solve some difficult philosophical problems.

Information, Prototypes, and Recognitional Capacities

The conception of recognitional capacities which I want to use stems from the one developed by Evans.[2] In elucidating the nature of recognitional capacities Evans uses the concept of information-based thought, which is defined as follows:

> a bit of information (with the content *Fx*) is in the controlling conception of a thought involving a subject's Idea of a particular object if and only if the subject's disposition to appreciate and evaluate thoughts involving this Idea as being about an *F* thing is a causal consequence of the subject's acquisition and retention of this information.[3]

Evans argued that a subject can be said to have a genuine thought about a particular object when making a judgement about a remembered object when he can *recognise* the object concerned. In his view:

> if a subject is disposed to identify a particular object as the object of his thought, and in so doing is exercising a genuine recognitional capacity stemming from the encounter or encounters from which the memory-information that saturates his thought derives, then . . . that object is the object of his thought, irrespective of whether it can be identified by means of any descriptions which the subject might otherwise use.[4]

If it is possible to give an account of information and recognitional capacities which does not make any philosophical presuppositions, then it might also be feasible that this could form the basis for a position which would address the genuinely philosophical problems of reference (to also deal fully with meaning would require a much richer account, including a more detailed explanation of mental representation).

One way in which the story might unfold is suggested by recent work on both real and artificial neural networks.[5] Information might be regarded as being present in the form of a pattern of activation, or vector, across a population of neurons. This is indeed how the brain represents sensory information such as smells, sounds and colours. These activation vectors undergo processing via vector-to-vector transformation; one population of neurons project via a vast number of synaptic connections to a second population, and through the careful tuning of those connections an appropriate activation pattern can be produced. Thus an incoming sensory vector can be instantly transformed into a prototype vector at higher populations of cortical neurons, perhaps via several transformations, and then onwards until finally an activation vector is produced in a population of motor neurons and behaviour is produced. We can thus view the pattern of synaptic connections in a network as the repository of information which has been gathered through interaction with the

---

[2] Evans, G. (1982) Ch. 8.
[3] Ibid. p. 122.
[4] Ibid. p. 269.
[5] For a description of the relevant literature, and its application to cognitive functioning, see Churchland, P. M (1995).

world. It is also possible to interpret the prototype vectors as the networks learned perceptual and explanatory categories. Appropriate behaviour is produced if sensory input is mapped onto the right prototype because these prototypes contain more information than is usually present in a given presentation, due to previous experience, and in the same way allow accurate expectations of other aspects of the situation.

This story is far from complete, and there are a number of difficult problems with it, but nevertheless it is highly suggestive; I will use at as a placeholder, for convenience of exposition, pending completion of a theory of this sort. We might view recognitional capacities as a type of vector prototype, as they take sensory information as input and produce an appropriate response as output when they are used "on-line". This neurocomputational background is needed in order to make sense of how a subject can be said to be exercising a genuine recognitional capacity when its purported object is not actually present, that is, when it is used "off-line". For if we had no independent means of access to that capacity then the account would be virtually empty, as it would suffer from the same problems as causal theories of reference with misrepresentation. The richness of the recognitional capacities account, compared to purely causal theories, comes from the more complex nature of recognitional capacities. A causal element is involved in the establishment of a recognitional capacity, but once in place that capacity can float free of its reference in a way which is crucial to the theory.

The Functioning of Recognitional Capacities

Recognitional Capacities for Individuals

An individual has a recognitional capacity for a particular object or person if they have a vector prototype, or some similar cognitive structure, which was formed through causal interaction with, and only with, that object. There may be a certain amount of vagueness concerning how long it takes for a recognitional capacity to become established, and this will depend upon the context, but once it is in place it allows us to attribute an ability to refer to that object. Where the recognitional capacity is brought about by contact with several indistinguishable objects the subject cannot be said to have a genuine recognitional capacity, and thus cannot be said to refer to a particular object. This allows the following definition:

> an individual, S, has a thought about an object, x, if and only if S has a genuine recognitional capacity for x, and that capacity is actively deployed in S's thinking.

This allows a definition of what it is for a term to be a proper name:

> a term, T, refers to an object, x, for an individual S, if and only if it is associated with a genuine recognitional capacity of S for x.

A similar approach could be taken towards demonstratives, given that the vector prototypes would be envisaged to be activated by both sensory and cognitive input. For demonstrative thoughts the activation would be caused by sensory input, and for cases where an object is being thought about without it being present the activation would originate from the internal cognitive environment. Hence:

>an individual, *S*, has a demonstrative thought about an object, *x*, if and only if the sensory
>input which *S* receives from *x* is mapped onto the appropriate vector prototype for *x*.

What the "appropriate vector prototype" amounts to needs additional specification, but it would involve mentioning previous encounters with *x*, or objects like *x*, if it has not been encountered before. Thus recognition of the *type* to which *x* belongs would be sufficient for attribution of a genuine demonstrative thought.

The problem of disjunctive meaning for cases of misrepresentation does not arise because of the complex role which the capacity plays within a human cognitive economy. For to credit the given cognitive structure as part of a recognitional capacities it must be "sensitive to considerations bearing upon the identity of a single object from time to time, and this means sensitive to spatio-temporal considerations".[6] Thus the person who possesses the alleged recognitional capacity must realise that appearance is a defeasible basis for re-identification, and must also have a conception of what kind of spatio-temporal information would be grounds for rejecting a supposed re-identification. It is important to note that the recognitional capacity is not *identified* with the vector prototype, as *individuals* have recognitional capacities. In other words "recognitional capacity" is a personal, not a sub-personal notion. The vector prototype is an important part of the grounding for the recognitional capacity, but the capacity involves many behavioural and cognitive sensitivities that need not be subserved by any one part of the cognitive system. Thus attributing a recognitional capacity to an individual does not commit one to exact claims about that individuals internal cognitive architecture. Such an entailment would be untenable, as it does not seem to be the case when we say of an individual that they can recognise *x*, that we are saying anything about their internal whirrings and grindings.[7]

One element of the sensitivity to spatio-temporal information would be an ability to keep track of position in space and time, and a related ability to restrict the domain of relevance for the recognitional capacity such that anything which falls outside, but which activates the capacity is not considered as a case of genuine recognition. If the subject were to be moved, without awareness of the move, to Twin Earth then they would produce judgements of re-identification which we can consider to be wrong. For were they to be informed of the move, they would withdraw their judgements, and it is this which allows ascription of error. The fact that such knowledge would not be available to the subject would not speak against the capacities status before the move. Thus for particulars the recognitional capacities account does yield determinacy of reference.

Recognition and Individuating Descriptions

It might be argued that nothing which has so far been said about recognitional capacities rules out the possibility that they might be, at base, a matter of possessing an identifying description. Evans produced an effective argument to show why this is not the case, and I wish to raise it briefly to show how it can be considered as compatible with the cognitive approach which I have sketched out above.

---

[6] Ibid. p. 278.

[7] See Wittgenstein, L. (1953) §149 for a statement of this point. I think that my own position (probably) does not succumb to Wittgenstein's argument.

Evans draws attention to a distinction between two ways in which we may exercise a memory ability, namely recognition and recall. It is a well established psychological finding that our ability to recognise information from a particular domain far outstrips our ability to recall information from that domain. It is obvious that in some sense one still has the information, but it is not available for the control of thought or action - it is not access conscious in Block's terminology[8] - and nothing which one only remembers in this thin sense can enter into the specification of thought content. So it is possible that one might have only the vaguest recollection as to what an individual one has met is like, yet still have a genuine capacity to recognise that individual when they are present. Thus we are not restricted, in recognising something, to observing whether or not it has features we are able to recall. This militates strongly against identifying the recognitional capacity with a description, given that any description which the individual will be able to formulate will be far from detailed enough to pick out the right object alone. The motivation for an account based on identifying descriptions comes from a mistaken view of recognition, which, as Evans points out, was diagnosed by Wittgenstein:

> It is easy to have a false picture of the process called "recognising"; as if recognising always consisted in comparing two impressions with one another. It is as if I carried a picture of an object with me and used it to perform an identification of an object as the one represented by the picture.[9]

Within the context of the present account it is worth noting that nothing like this comparison is postulated. It is the patterns of synaptic connections between neurons which ground a recognitional capacity, and these are not the sort of thing which could be accessible to consciousness. In addition, the process of vector transformation does not require the comparison of representations. Hence part of the motivation for using the notion of recognitional capacities is to avoid the need to find some descriptive means of capturing an individuals behaviour. For even if such a descriptive theory were possible for an individual it would not mean that the individual was using that theory in acting and judging. The desire to determine what somebody means by formulating such a theory is simply based upon a mistaken conception of meaning.

Recognitional Capacities and Natural Kinds

In the case of natural kind terms spatio-temporal considerations cannot be brought to bear to provide determinacy of reference, at least not in the same way as with particulars. I have argued that an account of meaning is exhausted when the behaviour of the relevant recognitional capacity is fully described. Yet this description need not necessarily solve the problem of what the recognitional capacity is a recognitional capacity *for*. Kripke and Putnam argue that the identity of the recognitional capacity should be given by the actual stuff which gave rise to the capacity in the first place.[10] The problem with this is what explaining "the actual stuff" amounts to. Kripke and Putnam

---

[8] See Block, N. (1995)

[9] Wittgenstein, L. (1953) §604.

[10] I shall ignore the differences between the two accounts, as I do not think that they matter in the present context. However, their views can be summarised as follows:

> **Kripke**: x is the natural kind instantiated by most or all of those items (which formed the initial sample).
>
> **Putnam**: x is ϕ if and only if x bears the relation same kind to the majority of the stuff I and my community have called "ϕ".

appeal to science to individuate natural kinds. They credit speakers with intentions to refer to a kind for which nature defines the boundaries, and they credit science with the ability to latch onto those natural kinds. I don't want to completely reject this picture, rather I want to argue that its status needs to be subtly readjusted.

The situation is confused by the need to distinguish between the public meaning of a term and the meaning with which it is imbued by ordinary speakers. This confusion is further exasperated by the ambiguity latent in the term "ordinary speaker". In dealing with these problems I will first describe how I think natural kind terms work for individuals in scientifically ignorant communities[11], and only then will I factor in the complexities brought about by the fact that contemporary speakers inhabit a community which encompasses a sub-group of scientific experts, and where the individuals of the community at large are more or less scientifically knowledgeable.

The crucial mistake which I think Kripke and Putnam make is to assume that natural kind terms, such as "water", have determinate reference even in scientifically ignorant communities. What is supposed to make their use determinate is an intention to refer to "that stuff" where the world sorts out what this amounts to, but it is my contention that the world cannot be appealed to in this fashion. The need for determinacy is important for Kripke and Putnam because in order to formulate an identity statement such as "water = H_2O" both terms have to be rigid designators, and thus must have determinate references. Treating "water" as though it has a determinate reference leads to externalism about meaning, because we can postulate molecular duplicates on Twin Earth for whom the reference of "water" is not H_2O but XYZ.

Prima facie postulation of determinate reference may seem like an acceptable move to make, but its advantages pale somewhat when we consider counterfactual scenarios. The sort of examples I have in mind are of the following kind: suppose that an alien race covertly replace large amounts of H_2O with XYZ on a planet inhabited by a scientifically ignorant civilisation (say earth before 1750) and that the change goes unnoticed. Given that we have authority to stipulate what happens in the possible worlds we describe, we can either have it that the community are unconcerned about the switch when told (after some time), or we can have them declare that what they *really* mean by "water" is the stuff which was around before the switch (remember that the identity statement "water = H_2O" is conceptually beyond them). Thus the problem with such examples is that what one wants to say about them seems largely to be a matter of intuition, backed by theoretical prejudice, and this is not a very principled way to resolve the issue.

The correct move here is to say that there is just no matter of fact about what they mean, or more accurately, there is no *exact* matter of fact. For they do not use the term "water" in a totally unprincipled way, rather an exhaustive specification of meaning simply remains silent about the philosophical question. We can describe how they use the term, and how they would use the term, given that it is grounded by a recognitional capacity. However, in order to get increased determinacy of meaning we would have to foist a scientific theory upon them, and that would amount to changing the hypothesised community in an illegitimate way. This is the case because the recognitional capacities which members of such a community possess are not sensitive to evidence about the fundamental properties of alleged referents, because they have no conception of such fundamental

---

[11] By "scientifically ignorant community" I mean a community which has *no* members whatsoever who are scientifically knowledgeable, or who engage in a science-like activity.

properties. In the face of this position the Kripke/Putnam line gains plausibility by citing instances where there are two substances which differ in fundamental properties (relative to our scientific theory) such as gold and iron pyrites (fool's gold). But in such cases the community treats them as different because there *are* methods available to the community for discriminating between them, and thus these can interact with what we might call the central aspect of the recognitional capacity to make members of the community able to refer to one or the other material. At the very least, in such situations a distinction in *public* meaning comes into being, a matter which I will discuss below.

A driving force behind the desire for determinate reference is realism. My position seems to be verificationist, in that it only acknowledges a determinate meaning where members of a community have a suitable means for detecting a difference between two proposed referents for a natural kind term. I would agree that my position is not compatible with what might be called full-blooded realism, which contends that there just are determinate primitive kinds in nature. However, I would also argue that such full-blooded realism is untenable. In some sense saying: "Water *really is* H_2O!" is right, but when this statement is uttered in a metaphysically loaded tone of voice it reveals a mistaken conception. The innocent reading refers to the possibility that there might be substances that behave like H_2O in a great many ways, in a fashion similar to the case of gold and iron pyrites, such that we can establish that this is not in fact the case. The full-blooded reading entails that in a completed science H_2O would constitute a primitive category such that it would be impossible for an alternative classification to arise which would allow us to categorise H_2O into different kinds. Thus essentially what the full-blooded view entails is that there has to a fundamental level in science in order for our natural kind terms to be properly grounded. Does the idea of a completed science make sense? How would we know that we had reached the most fundamental kind of particle, that quarks, or super-strings, or whatever did not themselves have components? Would such a science be genuinely explanatory, or would it have to postulate brute facts, such as charge or charm? I do not pretend to know what to say about these matters, but the important point for present purposes is that it is nonsense to hold that such considerations could be problematic for the meaning and reference of natural kind terms. Are natural kind terms so fragile? Must we live in terror of the day when a scientist discovers that H_2O is really a number of different substances?

I would answer no. The lesson from these considerations is that there is no such thing as a primitive relation between a community and a certain kind of substance. All kinds must be specified relative to a theory, or body of knowledge of some kind. This does not destroy the notion of a natural kind, rather it provides it with a firm foundation in the certain world of medium sized dry goods. More accurately, natural kinds lie at a level where we can make sense of differences between them, and so it is perfectly possible for this level to expand, indeed this is what has happened with scientific progress. But this just reflects the fact that meaning is relative to interests. If H_2O really did turn out to be composed of different substances, this would only affect usage if there was some practical difference between them, and would not constitute a semantic disaster. There is no absolutely correct thing to say in every imagined situation; it is wrong to ask: "Yes, but what do we *really* mean by 'water'?".

At every stage meaning would be exhaustively characterised by description of use, which includes the sensitivities of recognitional capacities to potential evidence, and so there is no threat of meaning being undermined by science, or of metaphysical meltdown. For on the conception I am

advocating there is no drive towards absolute determinacy at all points. Meaning is as it is, and if it leaves a certain style of philosophical questioning unsatisfied, then so much the worse for that style of questioning. If this is labelled verificationism, then I am willing to accept this consequence, so long as this does not prevent us from seeing the facts.

Due to the fact that, on this account, meaning is specified in terms that must be (potentially) available to the subject one can assert that meaning really is in the head, and that externalism about this type of meaning, recognition-based meaning, is mistaken. However, this does not mean that the externalist notion of meaning advocated by Kripke and Putnam is totally obsolete. Instead it should be seen as relevant to the specification of *public* meaning.

Recognitional Capacities and Public Meaning

The notion of meaning with which Kripke and Putnam are concerned is a term of art, and so does not feature in our ordinary talk of meaning. It constitutes something which we appeal to when we want (increased) determinacy of reference, and thus is maintained by the relevant group of experts in a community. So Putnam was right to talk of a division of linguistic labour, but wrong to view it as central to meaning. It is needed because there is a scale of complexity of recognitional capacities, from simple to complex. An example of a simple recognitional capacity is an infant's ability to recognise its mother. It is simple because the infant will respond to the right appearance-type regardless of whether it is actually its mother or not, there is no sensitivity to spatio-temporal evidence as the possible ground for rejecting a provisional re-identification. An example of a complex recognitional capacity is a physicist's ability to distinguish particle trails in the readout from a particle accelerator. What places a given recognitional capacity on this scale is the complexity of its interaction with potential forms of evidence.

Thus an individual can be credited with an intermediate recognitional capacity if they are aware that there is some body of expertise to which they could appeal if the wanted determinate reference for their terms. Such intentions to defer to experts need not be at the front of the subject's mind in every instance, but it must be the case that at *some* point they have *consciously* entertained the desire for their thoughts so to be grounded. It makes sense to attribute these sorts of intentions only in such cases, and not in instances where the subject has never had an awareness that there is a body of experts to which they could appeal. To attribute the intention that one's thoughts should be answerable to experts in such instances would be wrong, for how can we make sense of such attributions when the subject is not even capable of entertaining them?

The ambiguity which I alluded to above in the term "ordinary speaker" emerges here. For we attribute these sorts of intentions to members of ignorant communities (we treat them as though they were ordinary speakers of our community) in the course of producing a coherent interpretation of them, *by our lights.* However, it is important to be clear about the status of this explanation; it is something akin to taking the intentional stance towards them, and as such these attributions should be construed as purely pragmatic and instrumental. The explanation of such cases is dependent on the way we explain certain kinds of case for members of *our* community. For even if a member of our society is so ignorant that they have never had cause to entertain intentions to be well-grounded, we still treat them as answerable to public meaning. Evans illustrates how this relationship between an individual's utterances and public meaning works with the following example:

> A young student is reading out an ill-prepared essay to his class. It contains the sentence "A spark is produced electrically inside the carburettor". "That's not right", the teacher says. "What does he mean, class?" And here someone might say "He means the cylinder, sir". In saying this the second student is *not* committed to the idea that the subject had the thought, or even has the capacity to have the thought, "I shall say that a spark is produced in the cylinder". But nor is what he says independent of the subject's goals and beliefs. The point is rather something like this: to be saying that a spark is produced in the cylinder is what, given his general plans and situation, the subject should be doing; that is, doing that is what would conform best with the subject's plans at this moment.[12]

Now it is right to say that members of our (scientifically literate) community have this responsibility to public meaning. But, it is wrong to say that members of scientifically ignorant communities also have these kinds of intentions towards public meaning, because no such thing exists for them. In addition we cannot appeal to primitive relations to kinds-in-nature, so intentions to refer in the same way as at the initial dubbing are also illegitimate. Rather we extend this interpretive practice to them in order to make sense of their actions. We take their utterances as though they were made within the context of our speech community. This is acceptable as long as we do not take this to confer determinacy of reference upon their use of natural kind terms. Further, we should note that even individuals within our community do not have absolute determinacy of reference; their degree of accuracy is dependent on the present state of science. The crucial point here is to see that this does not lead to problems about meaning, as absolute determinacy is an unattainable ideal, which is not needed for meaning to play its proper role in our lives.

---

[12] Evans, G. (1982) p. 130.

Bibliography

Block, N. (1995) "A Confusion About a Function of Consciousness", *Behavioural and Brain Sciences*, **18**.

Churchland, P. M. (1995) *The Engine of Reason, the Seat of the Soul*. MIT Press.

Evans, G. (1982) *The Varieties of Reference*. Oxford: OUP.

Kripke, S. (1972) *Naming and Necessity.* Oxford: Blackwell.

Putnam, H. (1975) "The Meaning of 'Meaning'," in K. Gunderson, ed.*, Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press.

Wittgenstein, L. (1953) *Philosophical Investigations*. Oxford: Blackwell.