View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by DigitalCommons@Macalester College

Macalester Journal of Philosophy

Volume 15
Issue 1 *Spring 2006*

Article 2

6-1-2006

On How to Refer to Unobservable Entities

Gregory P. Taylor
Macalester College

Follow this and additional works at: http://digitalcommons.macalester.edu/philo

Recommended Citation

Taylor, Gregory P. (2005) "On How to Refer to Unobservable Entities," *Macalester Journal of Philosophy*: Vol. 15: Iss. 1, Article 2.
Available at: http://digitalcommons.macalester.edu/philo/vol15/iss1/2

This Article is brought to you for free and open access by the Philosophy Department at DigitalCommons@Macalester College. It has been accepted for inclusion in Macalester Journal of Philosophy by an authorized administrator of DigitalCommons@Macalester College. For more information, please contact scholarpub@macalester.edu.

On How to Refer to Unobservable Entities

Gregory P. Taylor

In order for us to associate a word with an object it might seem that we would need to have direct experience with both. Given the present technology, however, there are some objects with which we can have no direct experience, namely the unobservable entities postulated by scientific theories. The problem taken up here is how to refer to those entities. There are two prominent attempts to explain reference in scientific theories – the first is Ramsey and Carnap's proposal that we exchange theoretical terms for variables and existential quantification. The second is Kripke and Evan's causal theory of names and rigid designation. I will argue that the most plausible theory of reference to unobservables lies in between these two theories; terms that purport to refer to unobservable entities, when occurring within a theory, need to be thought of as bound variables. But when those same terms occur in sentences outside of the theory, as when spoken, for example, they occur as genuine referring expressions, which have their reference determined by a theory.

I. Theories

I will assume, for now, that it is legitimate to talk about ideal reconstructions of theories in which a theory is a set of assertoric sentences, to be distinguished from journal articles, textbooks, and the research program that revolves around the theory. A theory also needs to be distinguished from a scientist's expression *of* that theory. In what follows, "theory" should be taken to refer to that which both scientists and philosophers of science assume when they speak of the implications of a theory, of reducing one theory to another, or of a theory explaining a phenomenon. (As will become clear later, theories as such *must* exist in order for us to be able to refer to unobservable

entities.) Within the sentences that make up a theory, Carnap distinguishes between theoretical and observational sentences. Whereas observational sentences contain only terms that refer to objects of immediate experience, theoretical sentences contain reference to unobservable entities. The two are connected by correspondence rules that relate unobservable entities to observable ones. An example of such a rule would be: "The temperature of a gas is proportional to the mean kinetic energy of its molecules." Temperature is a property with which we are all familiar and which scientists can directly measure, whereas the only empirical evidence we can have for the mean kinetic energy of a set of molecules is through correspondence rules like this one. A theory consists of the theoretical and observational sentences plus the correspondence rules.[1]

II. Ramsey-Carnap Theory of Theoretical Terms

Imagine that we have a theory with the theoretical terms "electron," "proton," "atom," "mass," and "velocity" (among others left out for simplicity). Using Carnap's construal of a theory, we will have a series of theoretical sentences containing these terms, and correspondence sentences which contain these terms along with the observational terms O_1, O_2, O_3…O_n. The theory consists of a set of sentences, and can be represented in its entirety in the following manner:

(T)…electron…proton…atom…mass…velocity…O_1…O_2…O_3…O_n…

The "…" represent the connections between all of these terms (the other words), and "electron"…"O_3" represent all occurrences of those terms in the theory. We should actually imagine (T) as being the theory written out in full – this device has the sole purpose of making those terms stand out.

Ramsey realized that by turning all of the theoretical terms into variables and quantifying over the whole theory we get a new set of sentences that would have exactly the same predictive and explanatory power as (T).[2] Substituting "c_1" for every occurrence of "electron," "c_2"

---

[1] See Carnap, Rudolf *An Introduction to the Philosophy of Science* ed. Martin Gardner (New York: Basic Books 1966) 225-233

[2] See Ramsey, Frank "Theories" in *The Foundations of Mathematics* ed. R.B. Braithwaite (London: Routledge 1931), 221-236. The formulation that follows is parasitic on Carnap's formulation in *An Introduction to the Philosophy of Science*, 249-251

for "proton," "c_3" for "atom," "R_1" for "mass," "R_2" for "velocity," and then quantifying over the whole theory, we get:

(R) $\exists c_1 \exists c_2 \exists c_3 \exists R_1 \exists R_2 (\ldots c_1 \ldots c_2 \ldots c_3 \ldots R_1 \ldots R_2 \ldots O_1 \ldots O_2 \ldots O_3 \ldots)$.

This is called the *ramsey-sentence* for a theory. The variables, like the names in (T), are to be taken as classes of all particular occurrences of those variables in the theory, and what (R) does is replace each occurrence of a sentence like "mass(electron) = 10" in the theory with "there is a c_1 and an R_1 such that $R_1(c_1) = 10$," while maintaining the same structure and non-theoretical terms of the theory. Where (T) would say "electrons have properties x, y, and z," (R) says "there is a thing, and that thing has properties x, y, and z." The ramsey-sentence eliminates all occurrences of theoretical names, while maintaining the same assertoric force of the theory.

Quine would later use a similar device, in attempt to demonstrate the superfluity of singular terms altogether: instead of the sentences "Socrates is wise" we could say "there is a thing which is both Socrates and wise," thus converting Socrates from a name to a predicate, and eliminating reference.[3] Like Quine, Ramsey was interested in *eliminating* the problem of reference, and he proposed the ramsey-sentence as a means of avoiding using names.

Unlike Ramsey, though, Carnap was not content to leave things as they stand with (R); he wanted to *define* the reference of theoretical terms, not eliminate them. To accomplish this he introduced an *analytic-postulate*, which combines the standard expression of a theory, (T), and the ramsey-sentence, (R), and forms a conditional:

(A) $\exists c_1 \exists c_2 \exists c_3 \exists R_1 \exists R_2 (\ldots c_1 \ldots c_2 \ldots c_3 \ldots R_1 \ldots R_2 \ldots O_1 \ldots O_2 \ldots O_3 \ldots) \rightarrow$
(…electron…proton…atom…mass…velocity…O_1…O_2…O_3)[4]

This states that if (R) is true, then (T) is true too. What could make (R) true is a series of objects that *satisfy* (R). (The sentence "there is an x such that x is green" is satisfied by the objects *frog* and *leaf*, but not by the objects *White House* or *sun*.) If there is a series of objects that satisfy (R), then the corresponding theory that uses the theoretical terms "electron," "proton" etc… is also true.

---

[3] Quine, W.V.O. *Word and Object* (Cambridge: MIT Press, 1960), 179-187
[4] *An Introduction to the Philosophy of Science*, 271

Carnap's analytic-postulate fixes the reference for the said theoretical terms in the following manner: for (T) to be true the terms in it would have to refer ("electrons are x" can't be true unless there are electrons). If (R) is satisfied (and therefore true), (A) gives a stipulative definition of the theoretical terms, by telling us to replace each occurrence of "c_1" with "electron," "c_2" with "proton," and so forth. (A), then, fixes the reference of the theoretical terms. But since (A) *stipulates* the definition for these terms, "before [(A)] is laid down, these terms have no interpretation, not even a partial one. The only interpretation they receive in this form of the theory is the partial interpretation they obtain *through this A-postulate*."[5] This entails that if (R) is false (not satisfied) then no interpretation is given of the terms, because we learn nothing of the truth or falsity of (T).

(A) is an analytic postulate and all it tells us is that if there is a series of objects that satisfy the theory, then the terms in the theory refer to those objects. For the empirical content of the theory Carnap suggest that we use (R), which he calls the synthetic postulate. The whole picture, then, is this: (R) is the portion of the theory with empirical content – when we test the theory we are testing (R). What (A) tells us is that if (R) is true then the theoretical terms in (T) denote the objects that satisfy (R). (A) makes no empirical claims, because it is a *stipulative* definition.

Carnap's entire project here is an attempt to reformulate the analytic/synthetic distinction. Quine famously argued that when we test (T) we test *all* of (T), and therefore there is no distinguishing between those sentences which can be held true come what may (analytic) and those which can't (synthetic).[6] Carnap has found a holistic means of defining the terms of the theory so as to get around Quine's arguments (the terms are not defined individually, but the meaning of every term in the theory is stipulated at the same time, when (A) is laid down). (T) is logically equivalent to (A) and (R), but (A) has no empirical import (because it is *stipulative*).[7] And since (A) and (R) are logically independent of each-other, when we test the theory we test (R). Thus when we get a recalcitrant experience the adjustments will have to take place in (R); adjusting (A) could have no predictive consequences.

---

[5] *Ibid*, 271

[6] Quine, W.V. "Two Dogmas of Empiricism" in *From a Logical Point of View* (Cambridge: Harvard U. Press 1953) 20-46

[7] Quine himself admits in "Two Dogmas" that there is a clear distinction between the analytic and the synthetic in the case of stipulation, and this is precisely the "loophole" that Carnap takes advantage of; see *ibid,* 26

The new formulation of analyticity produces a new problem, though. Carnap's method can only interpret the terms if the theory is satisfied. All old theories become nonsense under this view, and as Philip Percival points out, "it is absurd… to suppose that 'electron' lacks denotation if current theory has the mass of an electron wrong at the second decimal place."[8] Perhaps we might try changing "true" to "mostly true." This doesn't get us anywhere, though. Even if we *assume* that our present theories are mostly true, and that most of the terms refer, we have absolutely no way of knowing which. For any term in the theory, it might be one of the few that don't refer. Also, what kind of a truth-value is mostly true? If the antecedent of a conditional is mostly true, does that make the consequent true or mostly true? Perhaps we could spell it out as saying that most of the sentences are true, but then, again, how are we to know which? The point here is not against scientific realism – our theories might be "mostly true." But whether or not that is the case, "mostly" truth won't allow an analytic postulate like (A) to fix the reference of theoretical terms, because mostly truth has no set implications for a conditional.

Ramsey argued that we could get along fine without theoretical terms; why did Carnap feel compelled to define them? He says something very peculiar about Ramsey's point:

> Ramsey certainly didn't mean… that physicists should abandon theoretical terms in their speech and writing. To do so would require enormously complicated statements…How can the sentence 'Mass (17) = 5' [where "17" is the name of a particular object] be translated into Ramsey's language? 'R_2 (17) = 5' obviously will not do; it is not even a sentence.[9]

Well, "R_2 (17) = 5" is not our only option; what about "$\exists R_2(R_2$ (17) = 5)"? Since 'Mass (17) = 5' is included in (T), "(R_2 (17) = 5)" will occur somewhere in (R) as a translation. Carnap rejects "R_2 (17) = 5" because it is an open sentence, but since it is contained in (R), then "$\exists R_2(R_2$ (17) = 5)" can be deduced from (R), just as we can deduce "$\exists x(Fx)$" from "$\exists x(Fx \ \& \ Gx)$." The reason, I think, why he doesn't consider this is because he infers, from the fact that R_2 is a variable, that the translation of "Mass(17) = 5" into the ramsey-sentence "must be supplemented by the assumptions concerning the relation R_2 that are specified in the ramsey-sentence."[10]

---

[8] Percival, Philip "Theoretical Terms: Meaning and Reference" in *A Companion to the Philosophy of Science*, 503

[9] *An Introduction to the Philosophy of Science*, 254

[10] *Ibid*, 254

The argument is that R_2 only has significance within (R) as a whole. But if we are convinced by Quine's holistic arguments, we know that this is also the case for "Mass." (Carnap must have found Quine convincing – otherwise he wouldn't have given a holistic reformulation of the analytic/synthetic distinction in the first place).[11] Without a theory behind it, the term "mass" doesn't have any more meaning than "R_2." Of course "$\exists R_2(R_2\ (17) = 5)$" will only have meaning in the context of the whole theory, but this doesn't mean that we can't assert it by itself, otherwise we wouldn't be able to assert "Mass (17) = 5" either; in both cases the sentence gets its meaning from the whole theory.

Carnap's account of theoretical terms (the (A)-postulate) was found wanting, because of its requirement that a theory be completely satisfied. But he had argued that we *must* use theoretical terms, and not variables, because open sentences tell us nothing and closed sentences can only be interpreted within the whole theory. This argument for the necessity of theoretical terms fails, because a regular theory, such as (T), is every bit as holistic as (R). If we can understand "Mass(17) = 5" then we can understand "$\exists R_2(R_2\ (17) = 5)$." Therefore Carnap's criticism of Ramsey, that the ramsey-sentence would require enormously complex statements, is unfounded, and since we could get along just fine with only the ramsey-sentence, Carnap's analytic postulate is unnecessary.

III. Causal Theory of Names and Natural Kinds

Two things were to be gathered from the last section: Carnap's theory of reference for theoretical terms failed, and we can get along just fine without it, using Ramsey's technique. But that it is possible to get along without theoretical terms doesn't entail that we actually do. Scientific theories don't seem to be either ramsey-sentences or nonsense, and so an account of theoretical terms is still needed. Fortunately, Carnap's theory of theoretical terms is not the only one available.

---

[11] He says of Quine and his new formulation: "Earlier, although I did not share the pessimism of Quine…, I always admitted that it was a serious problem and that I could not see a satisfactory solution… Finally, after many years of searching, I found this new approach, with the new A-postulate." *An Introduction to the Philosophy of Science*, 273-274. Carnap is not convinced by Quine's claim that we can't distinguish between analytic and synthetic, but rather that any attempt to do so must be both holistic and stipulative.

Kripke proposed the causal theory of names in the following passage from *Naming and Necessity*:

> Someone, let's say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of the chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman.[12]

When Feynman was born his parents named him and initiated a practice of using that name to refer to their son. Thus, if I say "Kripke told me that Richard Feynman was a physicist" I am participating in the social practice of referring to Richard Feynman (and Kripke). Regardless of what information I have about him, and regardless of whether or not that information is correct (suppose Kripke was mistaken when he said that Feynman was a physicist), I am referring to the person for whom the social practice was initiated. Indeed, we could only discover that what we said about a person was wrong if we actually referred to that person.

Kripke treats natural kinds the same way. Take gold, for example. "Suppose there were an optical illusion [due to the peculiar properties of the atmosphere] which made the substance appear to be yellow; but, in fact, once the peculiar properties were removed, we would see that it is actually blue."[13] Would we say in this case that gold didn't exist? Kripke says no – we would say that gold is actually blue. The word "gold" refers to that stuff, whatever it actually is, in the presence of which we commonly say "gold." (Think of Berkeley – he claimed that chairs are just ideas in our heads and not material objects; he did not say that chairs don't exist.)

Gareth Evans developed Kripke's brilliant insights into a full-fledged theory of reference. [14] As he sees it, the reference of a proper name is determined like this: a person is given a name (either at birth or later, as with a nickname) and there is a core group of people who frequently interact with her. They call her by that name, both in and out of her presence. Evans calls these people *producers*. Other people, who have never met her, may nonetheless pick up her name by hearing

---

[12] Kripke, Saul *Naming and Necessity* (Cambridge: Harvard U. Press 1972), 91

[13] *Ibid*, 118

[14] Evans, Gareth *The Varieties of Reference* ed. John McDowell (Oxford: Oxford U. Press 1982). Hilary Putnam did much the same thing, even anticipating Kripke on some points. See Putnam, Hilary "The Meaning of 'Meaning'" in *Mind, Language, Reality* (Cambridge: Cambridge U. Press 1975), 215-271

somebody use it or by being given a description about her (among other things). These people are called *consumers*. The essential condition for being a consumer is that one is not able to gather and spread any information about the person that was not already involved in the social practice of referring to that person; to do so is to become a producer.[15] The situation with natural kinds is almost identical.[16] Though just about everybody interacts with a certain species of plant, for example, they are not all necessarily producers. The real producers are those who can tell us what constitutes that plant, and distinguish it from other kinds of plants; we call these people "botanists."

It seems at first glance that this theory of names and natural kinds works perfectly for scientific language. Consider a term like "quark." Most people don't ever have anything to do (consciously) with a quark in their entire life. Nonetheless we are all able to refer to quarks. I just did, and the causal theory seems to provide an explanation of how it was possible. Scientists who actually interact with quarks publish articles about them, refer to them in class, and talk about them to their friends. All of the people who come into contact with the word "quark" in these ways are consumers, and they are thus able to refer to quarks, even though they (myself included) couldn't tell you the first thing about a quark. If asked what I mean by "quark," I would reply, "a particle or something; ask a physicist."

The causal theory does, indeed, provide a satisfactory explanation for most cases of reference. But as I will argue in the concluding section, closer examination of terms such as "quark" reveals an important difference between terms for observable and unobservable entities, a difference which necessitates two separate accounts of reference.

IV. The Reference of Theoretical Terms

To see the distinction, imagine that I've been hearing scratches in my walls and the pitter patter of little feet at night. I would infer that a mouse is living in my apartment. [17] Think of this as a scientific theory. I can go about testing it: I make the prediction that if I leave a piece of cheese on the floor when I go to bed it will be gone when I wake up. The situation here is similar to the situation with unobservable entities,

---

[15] *The Varieties of Reference*, 376-377

[16] *Ibid*, 382-383

[17] The example is from Van Fraasen, Bas C. "Concerning Scientific Realism" in *Philosophy of Science* ed. Martin Curd and J.A. Cover (New York: W. W. Norton & Company 1998), 1076

like quarks and electrons. Scientists make some observations and realize that there are some for which we can't account, given the present ontology, and so they postulate a new entity. When they want to go about testing their hypothesized entity they have to deduce, using the theory, further consequences that it will have for observation.

Suppose that I had named the elusive mouse Frank, and that I then proceeded to talk about him to my friends ("Frank kept me up all night"). The causal theory would hold that I am a producer for the practice of using "Frank" to refer to that mouse, and that the people who hear me talking about Frank are the consumers. If one of my friends were to say to another "I hope Greg catches Frank," they would be referring to the mouse in my apartment, even if they haven't had any contact with it.

The difference between the present example and a normal name, like Richard Feynman, is that here we have a name that isn't given an ostensive baptism, but rather is fixed by description. There is no previous object that we now call Frank; "Frank" refers to whichever object it is that is scratching, pitter pattering, and eating my cheese. There are many names that function like this: Jack-the-Ripper and Deep-throat, for example.[18] Whoever committed the infamous murders, and whoever was the Watergate informant, are the people to whom these names refer. This is exactly what is happening with an unobservable entity in a scientific theory: the theory is not saying "those things we call electrons have these properties," but rather, "there is a thing that has these properties – call it an electron." Before the theory comes along we have no knowledge of the properties or even the existence of *those things*. All unobservable entities are introduced this way – via description, not ostension. Otherwise they wouldn't be unobservable.

Every theoretical term must, at some point, have been introduced and had its reference fixed by a theory. We can then go about using that term to refer to the entity postulated by the theory. But the theory itself does not refer; it sets the rules for referring. According to Kripke, most people are given names in an initial "baptism" of sorts. When somebody says of their newborn "we will call him Jones" they are not *referring* to Jones, they refer to the boy with the word "him" and *name* him Jones – fixing the name, not using it. In the introduction of an unobservable entity, though, there is nothing to do the work that the

[18] *The Varieties of Reference*, 47-48. For names like this Evans uses what he calls the "the quantifier," as in: the person such that that person committed murders x, y, and z.

word "him" did in Jones' baptism. When the theory of quantum mechanics tells us about electrons it can not refer to electrons and then say something about it. The theory tells us what electrons are and tells us how to go about referring to them in the future – baptism by theory. If these considerations are correct we should actually think of the theory expressed by (T) as looking like this:

(TR) ∃ electron ∃ proton…(…electron … proton … O_1 … O_2 … O_3…),

which represents the implicit quantification of the name-fixing descriptions.

Concerning observable entities theories can refer just fine. When chemistry tells us that water is H_2O, we already know what water is. In this case the reference is not being fixed by a description. If we were to adopt a new theory of chemistry in which water is XYZ, we could still say that both theories were telling us the chemical composition of *water*, and therefore they are referring to the same thing. Kripke argued that even if everything we think about tigers turns out to be wrong, "tiger" will always refer to tigers (be they animals, machines, or holograms). This is crucial for Kripke's conception of a name; it is his concept of rigid designation – reference across possible worlds.[19] But it isn't clear that the same is the case with an unobservable entity. Take away everything we think electrons are and there is nothing left to point to and say "we were wrong about *that*."

The causal theory of names explains quite well how a lay-person can use theoretical terms. But whereas the producers for a normal object can perceive the object, the producers for an unobservable (the scientists) can only detect the effects of the object, through the use of a theory. This distinction is incredibly important, because it yields the distinction between ostension and description. We can point to a tiger and say "tiger," but before a theory has fixed the reference for the term "electron," we can only say "that which is causing the streak in the cloud chamber." ("That which," not "that.") Of course now we can say the word "electron," with meaning, but only because a theory has already determined that we should call that which causes the streak an electron.

Without the theory of quantum mechanics scientists simply have no interaction with electrons. (Unobservables are the noumena and the theory is the form that we bring to experience, so to speak.) It follows

[19] *Naming and Necessity* 48-49

that the theory, qua ramsey-sentence, must have existed *prior to* the beginning of the name-using practice. A name acquires its referent through the behavior of a group of producers, but because the scientists can only produce *after* the theory exists, the theory can not be participating in the name using practice. Theories do not refer to unobservable entities; unobservable entities satisfy theories, and then *we* can refer to those entities. Likewise, theoretical terms for unobservable entities refer to objects when spoken by a scientist or lay-person, but function as implicitly quantified bound variables when occurring within a theory.

Bibliography

Carnap, Rudolf. *An Introduction to the Philosophy of Science* ed. Martin Gardner (New York: Basic Books 1966)

Evans, Gareth. *The Varieties of Reference* ed. John McDowell (Oxford: Oxford U. Press 1982)

Kripke, Saul. *Naming and Necessity* (Cambridge: Harvard U. Press 1972)

Percival, Philip. "Theoretical Terms: Meaning and Reference" in *A Companion to the Philosophy of Science* ed. W.H. Newton Smith (Oxford: Blackwell 2000), 495-514

Putnam, Hilary. "The Meaning of 'Meaning'" in *Mind, Language, Reality* (Cambridge: Cambridge U. Press 1975), 215-271

Quine, W.V. "Two Dogmas of Empiricism" in *From a Logical Point of View* (Cambridge: Harvard U. Press 1953), 20-46

---. *Word and Object* (Cambridge: MIT Press, 1960)

Ramsey, Frank. "Theories" in *The Foundations of Mathematics* ed. R.B. Braithwaite (London: Routledge 1931), 221-236

Van Fraasen, Bas. "Concerning Scientific Realism" in *Philosophy of Science* ed. Martin Curd and J.A. Cover (New York: W. W. Norton & Company 1998) 1064-1087