Knowledge of Manifest Natural Kinds

Scott Soames

Manifest kinds are natural kinds designated by terms like—*water, tiger, gold, green,* and *electricity.* Individual instances of these kinds are objects of our potential acquaintance about which we may have *de re* knowledge. Natural kinds of a more highly theoretical sort—like photons and neutrons—are not included in this category. Manifest natural kinds, or *manifest kinds* for short, figure in interesting statements of theoretical identification, many of which are both necessary and knowable only aposteriori. The aim of this paper is to explain why this is so.

The statements I will be concerned with are expressed by sentences in which a manifest kind term combines with the copula to form a predicate applying to each instance of the kind. Examples of such predicates are *is water, is a tiger, is gold, is green, is electricity, and is hotter than.* Theoretical identification sentences containing such terms that express necessary aposteriori truths include the sentences in (1).[1]

1a. For all x, x is water iff x is H_2O. (Water is H_2O)
  b. For all x, if x is ice, then x is H_2O. (Ice is H_2O)
  c. For all x, if x is lightning, then x is electricity. (Lightning is electricity)
  d. For all material objects x, x is gold iff it is made up of the element with atomic number 79. (Gold is Au)
  e. For all x, if x is a tiger, then x is an animal. (Tigers are animals)
  f. For all x and y, x is hotter than y iff the mean molecular kinetic energy of x is greater than that of y.
  g. For all objects x, x is green iff x has surface spectral reflectance property SSR_{green}—the property of reflecting substantially more light in the middle-wavelength part of the visible spectrum than in the long-wavelength part, and approximately the same amount of light in the short-wavelength part as in the non-short part.

---

1 The parenthesized sentences—*Water is H_2O, Ice is H_2O, Lightning is electricity, Tigers are animals*, and *Gold is Au* have natural readings in which they are understood along the lines of the formulas preceding them.

In discussing these examples, I intend to focus on their most impor-
tant features, and to sidestep certain controversies surrounding them.
For example, some philosophers believe that in order for (1a) and (1b)
to be true, *is* in the predicate *is H_2O* must be understood as what is
sometimes called "the *is* of constitution," rather than as the normal
copula. Their thought is that instances of water are never instances of
H_2O; rather, instances of water are constituted by, or made up of, nu-
merically distinct instances of H_2O. For my own part, I don't think
there is any genuine contrast here, since I believe that just as an ice
cube can itself be an instance of the kind ice, while also being consti-
tuted by a numerically distinct instance of ice that is capable of sur-
viving the cube's destruction, so the ice cube may be an instance of
the kind H_2O, while also being constituted by a numerically distinct
instance of H_2O that is capable of surviving the melting of that ice.[2]
If this is right, then instances of ice and water not only are constituted
by H_2O, they are instances of H_2O themselves. However, this point
of controversy will not matter for my topic today. Whichever posi-
tion involving constitution and identity proves to be correct, all these
sentences are examples of the necessary aposteriori. Moreover, the
semantic nature of the predicates they contain plays an important role
in explaining why this is so.

The Positive Account

Let us first examine the way in which the necessity of these sentences
is related to their truth. As I see it, the crucial issue involves the
nondescriptionality of simple manifest kind terms, and the way in which
their reference is determined. What Kripke says about the general term
cat is the model for a great many terms for manifest natural kinds. He
says, "The original concept of cat is: *that kind of thing*, where the kind
can be identified by paradigmatic instances. It is not something picked
out by any qualitative dictionary definition."[3] Although this may be a
little cryptic, the point is clear. Just as ordinary proper names are
standardly introduced by stipulating that they are to apply to certain
objects with which one is already acquainted, so general terms like *gold,
water, tiger,* and *green* are standardly introduced with the intention that
they are to designate certain manifest kinds with which we are already

2   See Soames, *Beyond Rigidity*, (New York: Oxford University Press), 2002, chap-
      ter 11.
3   Page 122, *Naming and Necessity*, (Cambridge MA: Harvard University Press) 1980.

acquainted through their paradigmatic instances. For example, we may imagine these terms introduced by the following stipulations:[4]

> The general term *green* is to designate the color of all, or nearly all, paradigmatic *green*-samples (and none, or nearly none of the paradigmatic non-green-samples)—i.e. it is to designate the characteristic of object surfaces that is causally responsible for the fact that paradigmatic green-samples appear the same way to us (and different from paradigmatic *non-green*-samples). Hence, the predicate *is green* will apply (with respect to any world-state) to all and only those objects the surfaces of which have that characteristic which, in the actual state of the world, causally explains why the paradigmatic *green*-samples look the same to us (and look different from the paradigmatic *non-green*-samples.)[5]

> The general term *gold* is to designate the unique substance of which all, or nearly all, members of the class of paradigmatic *gold*-samples are instances (and of which none, or nearly none, of the class of paradigmatic *non-gold*-samples are instances). Substances are un-

---

4   Here I imagine the adjective *green* and the common nouns *gold, water,* and *tiger* being introduced with stipulations about what they are to refer to, or designate. The associated predicates—*is green, is gold, is water,* and *is a tiger* are then understood as applying to all and only instances of the kinds designated by the adjectives, or common nouns they contain. (For a more detailed account of the distinction between kind terms and the predicates formed by combining them with the copula see (i) my reply to Bernard Linsky in "Reply to Critics," in the forthcoming symposium on *Beyond Rigidity* in *Philosophical Studies*, and (ii) my reply to Gomez-Torrente in "Reply to Ezcurdia and Gomez-Torrente," in a similar symposium forthcoming in *Critica*.) Strictly speaking, the designata of the predicates, their extensions, are the sets of objects they apply to, rather than the kinds designated by the adjectives or nouns they contain. However, since the extension of one of these predicates (at a world-state) will just be the set of instances of the kind designated (at the world-state) by the adjective or noun it contains, the kind plays a crucial role in the semantics of the predicate. More on this later. Until then, it will do no harm to occasionally speak loosely of *the kinds designated by* these predicates.

5   It will be noticed that I allow the stipulation introducing the term *green* to make reference both to a set of paradigmatic positive instances (the items in the *green*-sample) and to a set of paradigmatic negative instances (the items in the *non-green*-sample). As indicated in chapter 7 of *Understanding Truth* (New York: Oxford University Press), 1999, this seems natural with color terms. How far this point extends to other terms, like *gold, water,* and *tiger,* is an open question. For that reason, I have formulated the stipulations introducing those terms so as to be compatible with this possibility. Thanks to David Manley for bringing this to my attention.

derstood to be physically constitutive kinds—i.e. kinds instances of which share the same basic physical constitution. Hence, the predicate *is gold* will apply (with respect to any world-state) to all and only those objects that share the basic physical constitution that nearly all the paradigmatic *gold*-samples actually have (and that none, or nearly none, of the paradigmatic *non-gold* samples actually have).

The general term *water* is to designate the kind instances of which share with all, or nearly all, members of the class of paradigmatic *water*-samples those properties that "make them what they are" (and that distinguish them from certain paradigmatic *non-water* samples). These are properties that explain their most salient characteristics— e.g. the fact that they boil and freeze at certain temperatures, that they are clear, potable, necessary to life, and so on. Hence, the predicate *is water* will apply (with respect to any world-state) to all and only those quantities of matter that have the properties that actually explain the salient features of all, or nearly all, of the paradigmatic *water*-samples (and that are lacking in all, or nearly all, the paradigmatic *non-water* samples).[6]

The general term *tiger* is to designate the species of animal of which all, or nearly all, the members of the class of paradigmatic *tiger*-samples are instances (and of which none, or nearly none, of the members of the class of paradigmatic *non-tiger*-samples are instances). Hence, the predicate *is a tiger* will apply (with respect to any world-state) to all and only those individuals that are members of the species of which nearly all the paradigmatic *tiger*-samples are actually members (and of which none, or nearly none, of the paradigmatic *non-tiger*-samples are actually members).

These stipulations are, of course, idealized. The manifest natural kind terms could have been introduced in this way, and they behave pretty much as they would if they had been so introduced. However, they need not have been introduced by any formal stipulation. It is enough if at some point speakers started calling relevant things *green, gold, water,* and *tigers,* with the intention that the predicates were to apply

6   This represents a slight change from the discussion of *water* in *Beyond Rigidity.* There, I treated it as governed by a stipulation analogous to the one for *gold* given above, rather than one mentioning the causal explanation of salient properties. See, however, pages 285–6 for a brief discussion of the possibility of treating *water* as standing for an explanatory kind of the sort indicated here.

not only to the particular objects speakers happened to encounter, but also to all and only instances of the relevant kinds to which those objects actually belonged. An analogous point holds for proper names. Although formal baptisms are common, there are also cases in which a proper name is introduced more informally, as when people start calling a certain body of water *Green Lake* and the habit catches on. In all of these cases, both the formal and the informal, we may speak of a manifest natural kind term or a proper name as being introduced ostensively.

When a manifest kind predicate P is introduced ostensively, and Q is another predicate related to P in a certain ways, there are often linguistic explanations of why the corresponding sentences with the logical forms (2a) and (2b) are necessary if true.

2a. $\forall x \, (Px \supset Qx)$
  b. $\forall x \, (Px \equiv Qx)$

I will consider two types of example. Case 1 is illustrated by the predicates *is gold* and *is a tiger*; case 2 is illustrated by *is water* and *is green*. In case 1, we let P be the ostensively introduced predicate *is gold* or *is a tiger*, and we let Q be a natural kind predicate that "designates" a kind of the same type as the kind "designated" by P—a substance, i. e. a physically constitutive kind, in the case of *is gold*, and a species of animal in the case of *is a tiger*. In speaking of the kind "designated" by a predicate, I mean to include two types of cases—those in which the kind **is** the semantic content of the general term from which the predicate is formed by adding the copula, and those in which the semantic content of the general term is property that determines the kind. Later I will say more about which predicates fall into which class. For now, suffice it to say that in both sorts of cases the extension, at an arbitrary world-state w, of a predicate that "designates" a kind k is the set of instances of k at w.

We now go through the reasoning, for case 1, of the necessity of true sentences of the forms (2a) and (2b): (i) From the assumption that the ostensive manifest kind predicate P has successfully been introduced it follows that there is a **unique** natural kind k_P of a given type T—a substance in the case of *is gold*, a species of animal in the case of *is a tiger*—of which nearly all members of the P-sample are instances (and nearly no members of the non-P sample are instances), and P applies, with respect to a world-state w, to all and only instances of k_P at w. (ii) From the assumption that the natural kind predicate Q also "designates" a natural kind, it follows that there is a kind k_Q which is such

that Q applies with respect to a world-state w to all and only members of k_Q at w. (iii) By hypothesis, the two predicates designate kinds of the same type; in the case of *is gold*, k_P and k_Q are both substances, while in the case of *is a tiger* they are both species of animal. (iv) Now suppose that (2a) is true. Since nearly all objects in the P-sample are P's, nearly all of the objects in the P-sample are also Q's, and hence instances of kind k_Q as well as kind k_P (v) But, by hypothesis, there is a **single** kind of the given type T—a single substance or species—of which nearly all members of the P-sample are instances (and nearly no members of the non-P-sample are instances). Assuming that we find that nearly no members of the non-P sample are B's, we may conclude that kind k_P = kind k_Q. (vi) This means that in addition to (2a), (2b) must also be true. (vii) Moreover, both must be **necessary**, since the extension of P with respect to a world-state w = the set of instances of k_P with respect to w = the set of instances of k_Q with respect to w = the extension of Q with respect to w.

Since sentence (1d) (about gold) is of the form (2b), it is necessary. (1e), which is about tigers and is of the form (2a), is slightly different. Here the term *tiger* is introduced with the stipulation that it is to designate whatever species of animal is the one of which paradigmatic members of the *tiger*-sample are instances (and of which paradigmatic members of the *non-tiger* sample are not). Thus, if the term is successfully introduced—in accord with the presupposition that there is one and only one species of animal satisfying this condition—then it will follow that the sentence *Tigers are animals* is true. It will even follow that it is a necessary truth, provided that it is an essential property of any species of animal that instances of it are themselves required to be animals. Given this, we conclude that (1e) is also necessary.

Next consider case 2, which is illustrated by the predicates *is water* and *is green*. As I have imagined their ostensive introduction, both stand for what might be called *explanatory kinds*. In the case of *is water*, the kind is one that is determined by the properties possessed by paradigmatic *water*-samples that both distinguish them from paradigmatic *non-water* samples and (causally) explain, in the actual state of the world, such salient characteristics of the members of the *water*-sample as their boiling point, freezing point, their properties as solvents, and so on. In the case of *is green*, the kind is one that is determined by the properties possessed by the surfaces of objects among the paradigmatic *green*-samples that distinguish them from paradigmatic *non-green*-samples and that actually explain why the members of the *green*-sample look alike to us, and different from members of the *non-green*-sample.

When P is a simple manifest kind predicate, like *is water* or *is green*, which stands for an explanatory kind, the necessity of (2a) and (2b) is accounted for a little differently than before. (i) We begin with the ostensive introduction of P by a stipulation that it is to apply (with respect to any possible state of the world) to all and only instances of the kind determined by certain properties—namely those possession of which by all, or nearly all, members of the P-sample in the actual state of the world distinguishes them from members of the non-P-sample, and (causally) explains the salient characteristics of the P-sample. (ii) It is then discovered scientifically that possession of the property expressed by Q distinguishes the members of the P-sample from the members of the non-P-sample, and (causally) explains the salient characteristics of the P-sample. (iii) From this it follows that the kind designated by the simple manifest kind predicate P is the kind determined by the property expressed by Q. (iv) This is sufficient to establish the necessity of sentences like (1a) and (1g) (about water and green objects respectively), which have the logical form (2b). Sentence (1f) (about heat and kinetic energy) is essentially the same, except for containing the two-place explanatory kind term *is hotter than*.[7] The explanation of the necessity of (1b), which is of the form (2a), is derivative from that of (1a).[8] (1c) (about electricity) is analogous to (1b).

In giving these linguistic explanations, I have appealed to the manifest natural kinds "designated" by different predicates without saying anything about what these kinds are. Although a number of fundamental questions about their metaphysical natures can be left open, some features of kinds are central to the linguistic model I have sketched. For one thing, the model presupposes that, whatever they are, manifest kinds are things that exist and have instances in different possible world-states. The color green, though not a green object itself, has green objects as instances. Moreover, since different objects are green in different possible world-states, the color green has different instances in different world-states. The color remains the same from state to state, even if all its instances vary. The same is true of the substances gold and water, as well as the species tiger.

In giving the model, I have also presupposed that manifest kinds are rather coarse grained, and are individuated by their possible instances in an interesting sense. If manifest kinds a and b have precisely the same instances in all possible world-states, then the kind a **is** the kind

7   See *Beyond Rigidity*, page 275.
8   See *Beyond Rigidity*, pages 294–97.

b; alternatively put, if two kinds are different, then they differ in at least some possible instances. Intuitively this seems plausible. It is hard to imagine two distinct species of animal, two distinct substances, or two distinct colors which have precisely the same instances in every possible world-state. The reason this is important for my linguistic model is that it is important that kinds **not** be individuated as finely as the properties that determine them. Consider, for example, the color green. Physicalists about color tell us that the object-color green is determined by a certain type of surface spectral reflectance property—one which specifies proportions of light reflected at different wavelengths.[9] Let Q be a complex predicate of English explicitly mentioning proportions of light reflected at different wavelengths that expresses this property. The predicate *is green* is clearly **not** synonymous with Q. The same can be said for other descriptive predicates. Suppose there is a further complex predicate Q' that applies to surfaces on the basis of a specification of their minute physical structure, which turns out to be necessarily equivalent to Q. Then, although the predicate *is green* is necessarily equivalent to both Q and Q', it is synonymous with neither. The different complex properties expressed by these predicates both have equal claim to determining the natural kind green, but neither is identical with the kind—that is the color—itself. What is said here about the predicate *is green*, and the color it designates, applies to linguistically simple manifest kind predicates generally, and the kinds they designate.[10]

The picture that emerges is one in which the meanings, or semantic contents, of linguistically simple manifest kind predicates like *is gold, is water, is a tiger*, and *is green* are complexes <I,K>, consisting of the meaning of the copula (roughly the relation of being an instance of) and the kind designated by the manifest kind term from which the predicate is constructed. Once this is seen, talk of the kind **designated** by such a predicate may be viewed as a loose approximation of something more precise. Normally in semantics, the designation of an expression is taken to be its extension, relative to a context and circumstance of evaluation. Since the extension of a predicate is the set of things to which it truly applies, simple manifest kind predicates **don't**, in this strict sense, designate kinds, though the simple manifest kind **terms** they contain do. Since, these terms are directly referential, the

9   See Alex Byrne and David R. Hilbert, "Colors and Reflectances," in Byrne and Hilbert, *Readings on Color, Volume 1: the Philosophy of Color* (Cambridge MA: MIT Press), 1997.
10 See pages 278–9 of *Beyond Rigidity*.

kinds designated are also their semantic contents. Thus, a simple predicate ⌜is T_S⌝ containing such a term **semantically expresses**, as its meaning, the complex <I,K> indicated above, in which K is designated by T_S. What the predicate designates (its extension), at any possible world-state, is the set of instances of K at that state.

By contrast, a manifest kind predicate ⌜is (a /the) T_C⌝ containing a complex general term T_C semantically expresses <I,D>, where D is the complex semantic content of T_C. When T is non-rigid, like *substance that falls from the sky and fills the lakes and rivers*, different kinds are designated by D (instances of which are in the extension of the predicate) at different world-states. When T_C is rigid, like *substance molecules of which have two hydrogen atoms and one oxygen atom*, the same kind is designated at every world-state. However, even when D determines the same kind in every world-state, it is not identical with that kind. Thus, when P is a simple manifest kind predicate the semantic content of which is <I,K>, and Q is a semantically complex predicate the semantic content of which is <I,D>, and D determines K with respect to every world-state, the sentences (2a) and (2b), though necessary, do not express the very same propositions as (3a) and (3b).

> 3a. $\forall x\ (Px \supset Px)$
>   b. $\forall x\ (Px \equiv Px)$

This point is crucial to blocking what would otherwise be a troubling argument that a number of the truths in (1) express propositions that are not only necessary, but also knowable apriori. However, blocking this argument is one thing, establishing that these propositions really are knowable only aposteriori is another. Example (1e), *Tigers are animals*, is particularly useful in illustrating the problem. Since I imagined the term *tiger* being introduced by a stipulation that it designate the animal species of which all or nearly all paradigmatic samples are instances, it may seem that it is part of the meaning of the predicate *is a tiger* that the things it applies to are animals, and, hence, that it is knowable apriori that tigers are animals. In my opinion, this is a mistake.

In order to see that it is a mistake, it is important to resist the temptation of an all-too-common line of argument. The tempting line goes something like this:

(i) To introduce a name or natural kind term n by stipulating that it is to stand for that which satisfies a certain condition, is to use a description D expressing that condition to semantically fix the reference of n.

(ii)   When D semantically fixes the reference of n, competent speakers associate D with n and know that the semantic rules governing n guarantee that it refers to whatever, if anything, satisfies D.

(iii) Because of this, competent speakers know apriori that which is expressed by **n is D** *(if there is a unique thing that is D)* when n is a proper name, and **For all x, x is an n iff x is an instance of the kind D** *(if there is such a thing as the kind D)* when n is a general term.

Here, we may let (ii) serve as a definition of what it is for a description to semantically fix the reference of a simple proper name or natural kind term. The idea it expresses is, essentially, the one behind Kripke's weak, fix-the-referent, version of the descriptivism defined by the first five theses listed at the beginning of lecture 2 of *Naming and Necessity*. Examples of expressions in the semantic literature that accord with the definition are descriptions rigidified using David Kaplan's *dthat*-operator.[11] Although the semantic content of **dthat D** is simply the denotation, if any, of D, in order to understand the rigidified description, one must know that its referent is, by definition, whatever satisfies D. The same is true of names and natural kind terms that have their reference fixed by descriptions in the sense of (ii).

Given this understanding of what it is for a description to semantically fix the referent of a term, we can isolate two mistakes in the reasoning from (i) to (iii). (i) is in error because it is possible to use a description as a tool to introduce a name or natural kind term without the description semantically fixing the referent of the term, and hence becoming part of what a competent speaker must master in order to understand it. For example, when looking at my first-born son and naming him *Greg Soames*, his mother and I did not intend the name to have the force of any *dthat*-rigidified description incorporating the content we used in singling him out. Although our stipulation relied on descriptive information to initially endow the name with meaning, that information was not incorporated into either the content of the name, or the conditions required to understand it. (iii) is in error because— whether or not a description used to introduce term also semantically fixes its reference — the description associated with the term does **not** give rise to apriori knowledge. In stipulating that *Greg Soames* was to be the name of our son, his mother and I did not thereby come to

---

11   David Kaplan, "Demonstratives," in J. Almog, J. Perry, and H. Wettstein, eds., (New York and Oxford: Oxford University Press), 1989.

know **apriori** that if we had a son, he was Greg Soames. On the contrary, it was because we were already acquainted with him, and knew him to be our son, that we were able to name him with our stipulation. The knowledge we relied on was, of course, aposteriori. After the name had been introduced, it was possible for us to express our knowledge in a new way—with the sentence *Greg Soames is our son*, and even more cautiously with the sentence *If we have a son, Greg Soames is our son*. But the knowledge expressed was just the old aposteriori knowledge we had before the name was introduced. The same can be said for all cases in which one uses a description to introduce a name, since, as I have argued elsewhere, it is a necessary condition on all such introductions that one believe, of the object to be named, that it is the denotation of the description used to introduce it.[12]

Applying these lessons to manifest kind terms leads one to several significant conclusions. First, although my imagined, idealized stipulations introducing *green*, *gold, water,* and *tiger* make use of descriptions of kinds in terms of certain paradigmatic instances (and non-instances), typically the descriptions do **not** semantically fix the reference of these terms. The descriptions usually don't enter into their meanings, and competence with the terms, and the predicates containing them, doesn't require speakers to understand anything about the paradigmatic instances (or non-instances) used to introduce them. Second, even those introducing the terms *green, gold, water,* and *tiger* don't know apriori of the particular objects mentioned in the introductions that they are green, gold, water, or tigers—any more than Greg's mother and I knew apriori that Greg Soames was our son. We did know, in virtue of knowing of our own stipulation, that he was named *Greg Soames* (if he indeed was our son). However, even that metalinguistic knowledge was not apriori, resting as it did on our knowledge of the empirical facts that endowed the name with its meaning.

By the same token, one who introduces the general term *gold* with the stipulation imagined thereby knows the metalinguistic truth that nearly all of the objects mentioned in the stipulation are ones to which the predicate *is gold* applies (if nearly all of them share the same physical constitution). However, such a person does **not** know apriori that nearly all those sample objects are gold, since that would require knowing that they are of the same physically constitutive kind, and that is

---

12 See chapter 16 of *The Age of Meaning: Volume 2 of Philosophical Analysis in the Twentieth Century*, (Princeton NJ: Princeton University Press), 2003.

something one can know only aposteriori.[13] Is anything weaker known
apriori? Does such a person know apriori that **if** nearly all of those
sample objects are of the same physically constitutive kind, then they
are gold? To ask this is to ask whether such a person knows apriori,
of the kind gold, that nearly all those sample objects are instances of
it, if nearly all of them are instances of any physically constitutive kind
at all.

In answering this question one must distinguish two related claims.

4a. The *gold*-stipulator knows apriori that if there is a physically
     constitutive k of which nearly all *gold*-samples are instances,
     then nearly all those samples are instances of k.
  b. If there is a physically constitutive kind k of which nearly all
     gold-samples are instances, then the *gold*-stipulator knows
     apriori that nearly all those samples are instances of k, if they
     are instances of any physically constitutive kind.

Although (4a) is trivially true, it has nothing to do with the semantics
of reference-determination for natural kind terms. Although (4b) is
relevant to the semantics of reference-determination, it is not true.
However, seeing this takes a little work.

There are two main cases to consider. In the first case, the stipulator
already knows aposteriori that nearly all the samples are instances of
the same physically constitutive kind, even though he is not able to
describe the kind in any very informative way except by reference to

---

13 Here I disagree with Kripke. See, for example, page 135 of *Naming and Necessity*,
    where he says the following: "the present view asserts, in the case of species terms
    as in that of proper names, that one should bear in mind the contrast between the
    a priori but perhaps contingent properties carried with a term, given by the way its
    reference was fixed, and the analytic (and hence necessary) properties a term may
    carry, given by its meaning. For species, as for proper names, the way the reference
    of a term is fixed should not be regarded as a synonym for the term. ... If we
    imagine a hypothetical (admittedly somewhat artificial) baptism of the substance
    [gold], we must imagine it picked out as by some such 'definition' as, 'Gold is the
    substance instantiated by the items over there, or at any rate, but almost all of
    them'. Several features of this baptism are worthy of note. First, the identity in the
    'definition' does not express a (completely) necessary truth: though each of these
    items is, indeed, essentially (necessarily) gold, gold might have existed even if the
    items did not. The definition does, however, express an *a priori* truth, in the same
    sense as (and with the same qualifications applied as) '1 meter = length of S': it *fixes
    a reference*." For a critique of Kripke's discussion of the apriori in connection with
    the meter stick see chapter 16 of *The Age of Meaning: Volume 2 of Philosophical
    Analysis in the Twentieth Century*. As to the parenthetically mentioned "qualifica-
    tions" mentioned in the passage just quoted, these are dealt with below.

the samples themselves.[14] It is, I think, reasonable to regard this as a case in which the agent is acquainted with the kind by virtue of being acquainted with some of its instances, and knowing of them that they are instances of a single physically constitutive kind. In this case, the agent's knowledge of the kind gold is aposteriori, and remains aposteriori when he introduces the predicate *is gold* to apply instances of it. The case is analogous to one in which I see one and only one man standing in front of me, and I think to myself, *He is standing in front of me.* My knowledge, of the man in question, that he is standing in front of me is based on, and justified by, my perceptual experience. Hence it is aposteriori. This fact would not change if I were to introduce the name *Saul* with the stipulation that it is to refer to the man standing in front of me. If I were to do that, the sentence *Saul is standing in front of me* would **not** express a proposition that I knew apriori. It would simply express a proposition that I already knew aposteriori, and that can be known only in that way. The case of the *gold*-stipulator who already knows that the items in his sample are instances of a single physically constitutive kind is similar. He already knows of the kind that his samples are instances of it. This knowledge is aposteriori, and remains so even after he has introduced the term *gold* to designate it.

But what about the knowledge mentioned in (4b)—knowledge of the kind gold that nearly all of the stipulator's paradigmatic samples are instances of it, **if they are instances of any one physically constitutive kind at all**? Isn't that something that the *gold*-stipulator knows apriori? No, it isn't. Think again about the man standing in front of me, of whom I know aposteriori that he is standing in front of me. My knowledge, of this man m, that if one and only one man is standing in front of me, then **he**, m, is standing in front of me is aposteriori, not apriori. Although this knowledge is based exclusively on the perceptual experience that presents the man to me, and hence allows me to grasp the proposition known to be true, my knowledge is also **justified** by that experience. As Jim Pryor has usefully reminded us, the fact that perceptual experience may play a crucial role in allowing me to entertain a certain proposition does not negate the fact that it may also play a crucial role in justifying my knowledge of that proposi-

14 The corresponding point for *green* and for *tiger*—namely that speakers may already be presumed to know, prior to introducing the term, that (nearly) all *green*-samples share some characteristic feature of their surfaces that explains their appearance and that (nearly) all *tiger*-samples are members of the same animal species—is quite plausible. Whether or not the same might be said for *gold* may be more controversial.

tion.[15] Hence, the knowledge I express by saying *If one and only one man is standing in front of me, then he* [demonstrating m] *is standing in front of me* is aposteriori, as is the knowledge I express by saying *If one and only one man is standing in front of me, then Saul is standing in front of me* [if I have introduced the name *Saul* to stand for that man]. The same is true of the *gold*-stipulator who says *If nearly all those samples are instances of a single physically constitutive kind, then nearly all of them are instances of it* [demonstrating the kind, gold], or *If nearly all those samples are instances of a single physically constitutive kind, then nearly all of them are gold* [if he has introduced the term *gold*]. What justifies this knowledge is that it is instances of a certain particular kind—gold—that he is acquainted with, and has empirically justified beliefs about. Hence, his knowledge is aposteriori, and for this *gold*-stipulator (4b) is false.

The second case to be considered is one in which we imagine the *gold*-stipulator as **not** knowing in advance that the items in the sample are of the same physically constitutive kind—even though in fact they are. The stipulation will be a little strange, if the stipulator doesn't at least **believe** that they are of the same kind, and take himself to have some evidence for this. However, if his evidence falls short of knowledge, then stipulatively introducing the term won't put him in a better epistemic position than before—any more than introducing the name *Saul* when I am not sure anyone is in front of me would improve my epistemic situation in that case. The *gold*-stipulator will, presumably, assent to the sentence *If nearly all these items in the sample are of the same physically constitutive kind, then they are all gold*, and to the extent that he is justified in believing that the sample does uniquely determine such a kind, he will be justified in believing the proposition expressed by the sentence.[16] He may even know this weaker proposition to be true. But if so, his knowledge is justified by the fact that the items he is perceptually acquainted with, and has beliefs about, **are** instances of one particular kind—gold—as opposed to any other. Hence, his knowledge is aposteriori, and (4b) is false.

As I see it, then, the situation is this. In order to successfully introduce a name or manifest kind term one must be acquainted with the object to be named or the natural kind to be designated. Standardly, this will involve being perceptually acquainted with, and believing certain things of, the object or the kind. In the case of manifest kinds, the

---

15 Jim Pryor, *Thinking about Water*, in preparation.
16 The same is true of the sentence *For all x, x is gold iff the predicate 'is gold' applies to x*, and the proposition it expresses.

normal way of being acquainted with, and believing things of, them is by being acquainted with, and believing things of, some of their instances. In order to successfully introduce a general term designating a manifest kind k, and to use it to express propositions of which k is a constituent, which one **knows** to be true, one must be acquainted with some particulars that are instances of k, and one must believe, or assume, with at least some justification that they are instances of a unique kind of the type that k is. In virtue of this, one is counted as knowing, aposteriori, of the particulars that they are instances of k, if they are instances of any relevant kind at all.

Having said this, I need to add three clarifying qualifications. First, I have **not** said that in order to have beliefs about the color green, the substances gold or water, or the species tiger, one must believe of some particular instances of these kinds that they are green, gold, water, or tigers (if they are instances of any relevant kind). Someone introducing these terms with the stipulations I have sketched must have such beliefs, but once the terms have been successfully introduced, they can be picked up by other competent speakers of the language, provided these speakers intend to use the terms with the semantic contents they have already acquired. These speakers need have no beliefs about particular instances of the kinds.

Second, some account must be given of what happens when a speaker introduces a term with a stipulation like the one I have given for *tiger*, say, without realizing that the supposedly paradigmatic items in the *tiger*-sample **don't** determine a species of animal at all. Borrowing from Putnam, we may imagine a world-state in which speakers stipulate that the predicate *is a tiger* is to apply to all members of the same species of animal as the tawny, stripped, cat-like individuals they have seen in various zoos, as well as in the wild—even though, unknown to them, these so-called tiger specimens are not animals at all, but cleverly disguised robots controlled by space aliens.[17] Putnam's intuition, which I share, is that in this fantastic scenario—in which speakers are under a monumental misimpression—the predicate *is a tiger* nevertheless turns out to be meaningful, and to truly apply to paradigmatic members of the *tiger*-sample. However, its meaning in the imagined world-state is **not** <I, K>, where I is the semantic content of the copula and K is the animal species which is the meaning of the term *tiger* for us, in the world as it actually is. The imagined world-state is **not** one in which tigers fail to be animals; there are no tigers in that

17 Hilary Putnam, "It ain't necessarily so," *Journal of Philosophy*, LIX, 1962, 658–71.

scenario (in our sense of *tiger*), even though there are things that speakers in the scenario correctly call *tigers*. Nor is the scenario one in which speakers wrongly **believe** that tigers are animals; speakers in the scenario have **no** beliefs about tigers, or the kind tiger, in our sense—even though they have beliefs which they express using the word *tiger.* Given all this, we can only conclude that there must be a process by which a word introduced with the intention that it is to designate a manifest kind of a certain sort may acquire quite a different meaning.

I suspect that what is going on is something like this: one who introduces the predicate *is a tiger* with the stipulation I have suggested intends (i) that it apply to nearly all specimens in the paradigmatic *tiger*-sample, (ii) that it apply to other things iff they bear a certain important relation of similarity to specimens in the sample, and (iii) that this similarity relation be the relation of being-an-instance-of-the-same-animal-species-as. In Putnam's fantastic scenario, these intentions cannot all be fulfilled, and the predicate acquires a different meaning by default—one which conforms to the first two intentions, but not the third. Depending on the beliefs and intentions of speaker-hearers, plus further empirical facts about the world-state, a new similarity relation comes to be the salient one—with the result that the predicate *is a tiger* acquires a meaning that is as close as is reasonably possible, given the situation, to the one intended by speakers.

The third point of clarification to be added to the picture presented above involves the question of whether it is part of the meaning of the predicate *is a tiger* that it applies only to animals. Nothing we have said so far settles this question. Since, in the Putnam scenario, the predicate means something other than what it actually means to us, the fact that it applies to non-animals in that scenario has no bearing on whether its actual meaning involves reference to animals.[18] Neither does the fact that *Tigers are animals* does not express an apriori truth. It is conceivable that it should be part of the meaning of a predicate P that it applies only individuals with the property expressed by F, even though the proposition expressed by ***P's are F*** is not apriori.

Consider the following analogy with names. Suppose I introduce the name *Philosopher-Saul* with the stipulation that it is to be synonymous the rigidified description *dthat [the x: x is a philosopher and x =*

---

18 By the same token, Putnam's scenario is essentially irrelevant to the question of whether the propositions we actually use sentences containing the word *tiger* to express are knowable apriori. If we hadn't already concluded that it is not knowable apriori that tigers are animals, Putnam's scenario would not justify drawing this conclusion.

Saul Kripke]. In order to understand the name, one must know that it refers to an individual iff that individual is both Saul Kripke and a philosopher.[19] Hence, it is part of the meaning of the name that it refers to a philosopher, if it refers to anything at all. However, the proposition semantically expressed by

 5. Philosopher-Saul is a philosopher (if *Philosopher-Saul* has a referent).

is just the singular proposition that says of Saul Kripke that he is a philosopher (if the name *Philosopher-Saul* has a referent). Since this proposition can be known only aposteriori, it is **not** (strictly speaking) knowable apriori that Philosopher-Saul is a philosopher (if *Philosopher-Saul* has a referent). This is true, even though the aposteriori knowledge needed to understand (5) is, arguably, sufficient for knowledge of the proposition (5) expresses. This illustrates the larger point that in order for a proposition p to be knowable apriori, it is not enough that there be some sentence S which both expresses p and is such that understanding S provides one with all the justification one needs to know p. Such sentences and propositions do have an interesting epistemological status; the sentences might well be termed *analytic*, and the propositions they express can be known to be true without any empirical justification beyond that required to understanding sentences that express them. Nevertheless, these propositions are not knowable apriori.[20]

Applying this lesson to the predicate *is a tiger*, we get the result that **if** it is part of the meaning of the predicate that it applies only to animals, then, even though the proposition that tigers are animals is **not** knowable apriori, it **is** knowable solely by virtue of the knowledge needed to understand the (analytic) sentence *Tigers are animals*, which expresses it. Whether or not this is part of the meaning of the predicate is a question that I leave open.[21]

19 It is worth noting that this shows that character in David Kaplan's sense—a function from contexts of utterance to contents—cannot, in general, be identified with the meaning of an expression, in the sense of that knowledge of which is necessary and sufficient for understanding the expression. If meaning were identified with character, then in a language containing *dthat*-rigidified descriptions, **dthat *[D₁]*** and **dthat *[D₂]*** would be synonymous whenever D_1 and D_2 were necessarily co-designative, and grasping the contents of the unrigidified descriptions would not be required for understanding the rigidified descriptions. These results are clearly unacceptable.

20 For further discussion, see chapter 16 of *The Age of Meaning: Volume 2 of Philosophical Analysis in the Twentieth Century*.

21 For further discussion see chapter 4 of *Reference and Description: The Case Against Two-Dimensionalism*, (Princeton and Oxford: Princeton University Press), 2005.

This completes my explanation of the necessity and aposteriority of sentences containing simple manifest kind predicates like those in (1). As I see it, the necessity of many of these statements follows from their truth, plus the way in which the reference of the terms they contain is standardly fixed. The explanation of their aposteriority is based on the idea that our knowledge of manifest kinds parallels our knowledge of individuals. Just as our *de re* knowledge of individuals standardly depends either on our own acquaintance with them, or on the acquaintance of others who pass important parts of their knowledge on to us, so our *de re* knowledge of manifest kinds standardly depends either on own acquaintance with members of these kinds, or on the acquaintance of others who pass aspects of their knowledge on to us. Because of this requirement on acquaintance, most of our knowledge of individuals, and of manifest kinds, is aposteriori. It is not possible to circumvent this result by using descriptions to introduce or to semantically fix the reference of names or manifest kind terms. In both cases, the requirement that we antecedently believe of the object to be named, or the kind to be designated, that it is denoted by the description used to introduce the term renders our knowledge of the propositions expressed by relevant sentences containing the term aposteriori, rather than apriori.[22]

Contrast with Two-Dimensionalism

The account I have offered of natural kind predicates for manifest kinds is both Millian and nondescriptional. It is Millian in that it holds that the semantic contents of the general terms out of which simple manifest kind predicates are constructed are the kinds they designate. It is nondescriptional in recognizing that although descriptions may be used to introduce these general terms, the terms themselves are standardly not synonymous with rigidified descriptions. Moreover, even when descriptions are involved, their use does not give rise to apriori knowledge, but rather must be grounded in empirical, *de re* knowledge of their denotations. These features of my account stand in marked contrast to competing two-dimensionalist accounts, all of which treat simple general terms like *green, gold, water,* and *tiger* as fundamentally descriptive, in one way or another. Since my space is limited, I will here say

22 One can, of course, know apriori that for all x if x is gold, then x is gold, even though simply entertaining the proposition involves some kind of acquaintance with the kind, which in turn may require some empirical knowledge. The reason such empirical grounding does not prevent knowledge of this proposition from being apriori is that it plays no role in justifying the proposition we apprehend.

just a word about one particular type of two-dimensionalist account—something I call *weak two-dimensionalism*.[23] Its central tenets are:

Tenets of Weak Two Dimensionalism

T1. Each sentence is semantically associated with a pair of semantic values—primary intension, and secondary intension. The primary intension of S is its Kaplan-style character. The secondary intension of (or proposition expressed by) S at a context C is the proposition assigned by its primary intension to C.

T2. Understanding S consists in knowing its primary intension—i.e. its meaning, or character. Although, this knowledge, plus complete knowledge of the context C, would give one knowledge of the proposition expressed by S in C, one often does not have such knowledge of C. Since we never know all there is to know about the designated world-state of C, sometimes we don't know precisely which proposition is expressed by S in C. However, this does not stop us from correctly using S in C.

T3a. Examples of the necessary aposteriori are sentences the secondary intensions of which are necessary, and the primary intensions of which assign false propositions to some contexts.

T3b. Examples of the contingent apriori are sentences the secondary intensions of which are contingent, and the primary intensions of which assign true propositions to every context.

T4a. All proper names and natural kind terms have their reference semantically fixed by purely descriptive properties, which can, in principle (given a rich enough vocabulary), be expressed by descriptions not containing proper names or natural kind terms.

T4b. These names and natural kind terms are synonymous with descriptions rigidified using *actually* or *dthat*.[24]

T5a. **It is a necessary truth that S** is true w.r.t. a context C iff the secondary intension of S in C is true w.r.t. all world-states that are possible relative to C.

23 For a thorough critique of all major forms of two dimensionalism, see my *Reference and Description*.

24 The character of **dthat [the D]** is a function from contexts to the denotation o of **the D** in the context; propositions expressed by sentences containing **dthat [the D]** are singular propositions about o. The character of **the x: actually Dx** is a function from contexts C to the property of being the unique object which "is D" in C_W (the world of C); propositions expressed by sentences containing the description are singular propositions about C_W.

T5b. Standardly, an attitude ascription *x v's that S*, when taken in a
       context C, is true of an agent i iff there is some meaning (charac-
       ter) M such that (i) i bears R to M, and (ii) M assigns to i's con-
       text the **secondary intension** of S relative to C. So propositions
       are objects of the attitudes, and attitude verbs are two-place predi-
       cates of agents and their objects. However, this two-place rela-
       tion holds between an agent a and a proposition p in virtue of a
       three-place relation holding between a, a character, and p. To
       believe p is to accept a character M that expresses p (and to be-
       lieve that M expresses a truth). To know a true proposition p is
       to justifiably accept a character M that expresses p (and to know
       that M expresses a truth).

T6a   For all propositions p, p is both necessary and knowable only
       aposteriori iff (i) p is necessary, (ii) p is knowable **in virtue of**
       one's justifiably accepting some meaning M (and knowing that it
       expresses a truth), where M is such that (a) it assigns p to one's
       context, (b) it assigns a false proposition to some other context,
       and (c) one's justification for accepting M (and believing it to
       express a truth) requires one to possess empirical evidence, and
       (iii) p is knowable **only** in this way.

T6b   For all propositions p, p is both contingently true and knowable
       apriori iff in addition to being contingently true, p is knowable **in
       virtue of** one's justifiably accepting some meaning M (and know-
       ing that it expresses a truth), where M is such that (a) it assigns p
       to one's context, (b) it assigns a true proposition to every context,
       and (c) one's justification for accepting M (and believing it to
       express a truth) does not require one to possess empirical evidence.

I will close by using terms for manifest kinds to highlight certain prob-
lems with weak two-dimensionalist theories of this sort. As is indi-
cated in thesis T4a, a crucial feature of these theories is their analysis of
natural kind terms as having their reference semantically fixed by de-
scriptions. For example, it is common for two-dimensionalists to main-
tain that the word *water* has its reference semantically fixed by a de-
scription expressing widely shared knowledge about water—something
like the description *the clear, potable stuff that fills the lakes and oceans.*[25]
Accordingly, these two-dimensionalists take *water* to be synonymous

---

25 See, for example, David Chalmers, *The Conscious Mind*, (New York and Oxford:
    Oxford University Press) 1996, p. 57 and Frank Jackson, "Reference and Descrip-
    tion Revisited," *Philosophical Perspectives* 12, 1998, p. 212.

with a rigidified version of this description, and they take the proposition expressed by

> 6.   Something is water iff it is an instance of the clear, potable stuff that fills the lakes and oceans, …

to be an example of the contingent apriori. On my view, this is incorrect. Since the proposition expressed by (6) contains the kind water as a constituent, knowledge that it is true requires *de re* knowledge of the kind, which cannot be had purely descriptively, but rather requires grounding in aposteriori knowledge of some instances of water.[26]

The two-dimensionalist's failure to see this is, I think, rooted in the unwavering descriptivism expressed by T4a and T4b. Both are problematic. Contrary to T4a, it is very hard—I believe impossible—to come up with any plausible description to play the role of a semantic reference-fixer for a manifest kind term like *water*, just as it is very hard to come up with such a reference-fixer for a name like *Bill Clinton*. The reason is the same in both cases: although virtually everyone who uses the terms will associate them with some descriptive information, that information may vary widely from speaker to speaker, and none of it is required in order for someone to understand the terms. Rip Van Winkle, who knew Clinton by name when Clinton was a pre-teen and who wakes up today knowing only that he has slept for a long time, understands the sentence *Bill Clinton is not very honest* perfectly well, and he surely counts as a competent user of the name, even though there is little overlap between the information he associates with it and the information we do. The same is true of the term *water*. Imagine someone who has lived in a very restricted environment—someone who has been confined to a basement, who has never been outside, who knows nothing of oceans or lakes, who has heard the word *water* and knows that it applies to cloudy stuff running through a drain in his basement from a nearby laundry, but doesn't know that water is normally clear or drinkable. Such a person may nevertheless correctly use the word *water* to refer to water, just as we do, and he may understand many sentences containing it, despite not knowing their truth values. What is important is that he has been in contact with the stuff and knows that the word applies to it—just as with Rip Van Winkle and Clinton.

Observations like these cast doubt on T4a. However, that thesis is not the only problem for weak two-dimensionalism. Even if one could

---

26 Jim Pryor develops a similar critique of two-dimensionalism in his manuscript *Thinking about Water*.

find plausible candidates for descriptions semantically associated with natural kind terms, the rigidification required by T4b would itself be problematic. Here, weak two-dimensionalists face a dilemma. On the one hand, if they use the actuality operator to rigidify, and analyze the secondary intension of *Water is H₂0* as equivalent to that expressed by

7.   Instances of the kind which is actually D are instances of H_2O

then, they will wrongly characterize the truth of attitude ascriptions like

8.   Even if it had been the case that so and so, John would have believed that water was H_2O

as requiring agents like John in other, merely possible, world-states to have beliefs about the actual world-state in which (8) is assertively uttered.[27] On the other hand, if they rigidify using the *dthat* operator, and analyze the secondary intension of *Water is H₂O* as the proposition expressed by

9.   Instances of the kind dthat [D] are instances of H_2O

then, that proposition will contain the kind water as a constituent—in which case, knowledge of its truth must be grounded in some acquaintance with the kind water. Since such acquaintance is **not** required (for arbitrary D) in order to accept the meaning of (9) and know that it expresses a truth, thesis T5b's account of what it takes to know that water is H_2O is liable to be incorrect, as is the general account of the necessary aposteriori given in T6a. In short, the weak two-dimensionalist theory faces apparently intractable difficulties no matter which form of rigidification is chosen.[28]

So much for my brief critique of the weak two-dimensionalist treatment of manifest kind terms. It is worth noting that this last problem, involving the rigidification required by T4b, does not arise for versions of what is probably a more familiar form of two-dimensionalism—one which I call *strong two-dimensionalism*. The chief distinguishing characteristic of this view is its rejection of the analysis of attitude ascriptions given in T5b in favor of a more radical thesis that takes belief and knowledge ascriptions—including those involving operators like *it is knowable apriori that* and *it is knowable only aposteriori that*—as operating on the primary intensions, rather than the secondary intensions, of

27 This is a straightforward extension to natural kind terms of the argument found in chapter 2 of *Beyond Rigidity*, pp. 39–50.

28 These (and other) arguments are developed at much greater length in chapter 10 of *Reference and Description*.

their complement clauses.[29] According to this view, what is reported by the knowledge ascription *x knows that water is H₂O* is **not** that the agent knows the secondary intension of *Water is H₂O*, but rather that the agent knows its primary intension—something he will know just in case he knows of the meaning of *Water is H₂O* that it expresses a truth. Since rigidifying operators in the complement clauses of these knowledge ascriptions make no significant contributions to their primary intensions, none of the problems with T4b carry over to versions of strong two-dimensionalism.

In the end, however, this is no help, since, as I have argued elsewhere, strong two-dimensionalism is independently refutable on other grounds.[30] Although this is a bad result for two-dimensionalist semantic theories, it does not affect our ability to explain instances of the necessary aposteriori involving natural kind terms that designate manifest kinds. As I have argued above, by adhering to the fundamentals of Kripke's nondescriptionalist account of these terms we can explain both the necessity and the apriority of examples like those in (1), without making unrealistic semantic or metaphysical assumptions.[31]

Scott Soames
School of Philosophy
University of Southern California
3709 Trousdale Parkway
Los Angeles, CA 90089-0451
USA
soames@Princeton.EDU

29 There are several other differences between strong and weak two-dimensionalism that accompany this one—including a somewhat different characterization of primary intension (which must, on the strong two-dimensionalist view, be a proposition). In addition, strong two-dimensionalists are often inclined to take propositions to be sets of possible world-states. For a full discussion see *Reference and Description.*

30 See chapter 10 of *Reference and Description*, as well as my "Kripke, the Necessary Aposteriori, and the Two-Dimensionalist Heresy," in M. Garcia-Carpintero and J. Macia, eds., *The Two-Dimensional Framework: Foundations and Applications*, (Oxford University Press), forthcoming.

31 This paper is a slightly updated version of a talk given on June 5, 2003 at the Third Barcelona Workshop on Issues in the Theory of Reference. The general point of view it presents—especially the critique of two-dimensionalism—is developed more fully in my forthcoming book, *Reference and Description* (2005). As for the positive view of natural kind terms developed here, a substantial portion of the middle part of this paper—on natural kind terms, reference-fixing, and the apriori—has been incorporated into a section, with that title, of chapter 4 of the book.