1

Essential Properties, Thought-Experiments, and Modal Intuitions

I WANT to open these explorations by considering Kripke's account of *essential properties*. My aim is not only to understand what sort of account this is, i.e., how Kripke arrives at and argues for judgments regarding accidental and essential properties, but also to examine how one might engage critically an account of the sort Kripke advances.

I begin the story in an epistemological tone of voice. The claim

(1) Gold is yellow,

Kripke notes, has historically been advanced as an example of an *a priori* truth, something we know "prior to" and "independently of" specific experience. Kripke is initially concerned to challenge this purported epistemic status and to establish that, if (1) is in fact true, its truth is nevertheless not something we know *a priori*. To this end, he proposes a thought-experiment:

> Suppose there were an optical illusion which made the substance appear to be yellow; but, in fact, once the peculiar properties of the atmosphere were removed, we would see that it is actually blue. Maybe a demon even corrupted the vision of all those entering the gold mines . . . , and thus made them believe that this substance was yellow, though it is not. (118)

Here, according to Kripke, we have the description of a situation in which, contrary to our present beliefs, (*a*) gold is not yellow and, indeed, (*b*) gold never was yellow. If we were to discover tomorrow that atmospheric conditions or corrupting demons have continuously been producing the envisioned effects, claims Kripke, we would not conclude that there is no gold or that what we took to be gold was not in fact gold. Rather, we would conclude that gold only *appears* (and always has appeared) to be yellow, but that it has turned out that gold is actually blue. We would conclude, that is, that (1) expresses what has proved to be a *mistaken belief* about the color of gold. Thus, Kripke concludes, even if gold really is yellow, as it appears to be and as we now believe it to be, *that* gold is yellow is not something we know *a priori*. Being yellow is not "part of the concept" of gold. Our present belief (1) rather derives from our various experiential encounters with gold, and further such experiences and matter-of-factual discoveries could show that belief to be false; (1) is empirically defeasible.

We will return to Kripke's account of (1) shortly. First, however, I want to begin the exploration of the new instruments Kripke introduces in "Naming and Necessity" for dealing with traditional questions of modality and identity: the idea of a rigid designator, a strategy of argument by counterfactual stipulation, and a causal-historical account of the reference of proper names and other designators.

What linked these three instruments into a functioning unit was the further centripetal notion of a *possible world*. Kripke introduces the notion of a rigid designator, for example, by telling us that a rigid designator is an expression which designates the same object "in every possible world"(48). A possible world, in turn, is "given" by a counterfactual description.

> A possible world is *given by the descriptive conditions we associate with it.* (44)

> 'Possible worlds' are *stipulated*, not *discovered* by powerful telescopes. (44)

There are, however, constraints on those descriptive specifications suitable for "stipulating" possible worlds. Not just any subjunctive supposition will do. A *sine qua non* condition is that such a stipulation be framed in our current language "with *our* meanings and *our* references".

> When I say that a designator is rigid, and designates the same thing in all possible worlds, I mean that, as used in *our* language it stands for that thing, when *we* talk about counterfactual situations. I don't mean, of course, that there mightn't be counterfactual situations in which the other possible world's people actually spoke a different language. One doesn't say that 'two plus two equals four' is contingent because people might have spoken a language in which 'two plus two equals four' meant that seven is even. . . . [We may even be] describing a possible world or counterfactual situation in which people, including ourselves, did speak in a way different from the way we speak. But still, in describing that world, we use *English* with *our* meanings and our references. It is in this sense that I speak of a rigid designator as having the same reference in all possible worlds. (77–78)

What these remarks imply is that, when "stipulating" a possible world, we need to hold fixed the references of any *rigid* designators we use in describing that world. Correlatively, then, our confidence in the conclusions we propose to draw from consideration of such a counterfactual stipulation extends no further than our ability to conform to this constraint, and so, *inter alia*, our ability to recognize which of the designators we have used in that stipulation are in fact rigid designators.

But now it may well seem that we are traveling in a circle, for to identify these rigid designators, we will need to determine which of the expressions we have used in the stipulative specification of *that* possible world in fact designate the same object in *every* possible world. In other words, we have so far characterized rigid

designators in terms of possible worlds, possible worlds in terms of admissible counterfactual descriptions, and the admissibility of a counterfactual description in terms of rigid designators. This may well provoke some echoes.

Recall, for example, Quine's "Two Dogmas of Empiricism". Analyticity was to be explained as truth by virtue of meaning, more specifically, as derivability from logical truths upon substitution of synonyms for synonyms. Synonymy, in turn, was to be explained in terms of substitutivity *salve veritate* in all contexts, including modal ones, and specifically in contexts of necessity. And necessity, of course, was to be explained in terms of the analyticity of certain biconditionals. The critical movement in "Two Dogmas", in other words, was similarly circular: from analyticity to synonymy, from synonymy to necessity, and from necessity back once again to analyticity.

As Quine recognized, what one needed to break out of this circle was a bit of epistemology. In particular, one needed a non-arbitrary strategy for *establishing* synonymies. One needed for there to be a fact of the matter concerning the synonymy or non-synonymy of two expressions (whether in the same or in different languages), and so for there to be a *method of inquiry* for determining in specific cases what that fact of the matter in fact was. Quine, as is well known, claimed that we couldn't have the requisite bit of epistemology, at least that there could be no non-arbitrary *empiricist* strategy for establishing term synonymies—hence the indeterminacy of translation, the inscrutability of reference, ontological relativity, and all that. I've argued elsewhere[1] that Quine was wrong about this, but that isn't currently to the point. What *is* now to the point is that, in order to break out of Kripke's circle of elucidations, we might suppose, one also needs a bit of epistemology. It is puzzling, therefore, to discover that, at the crucial moment, Kripke explicitly rules out the relevance of epistemological considerations to questions of contingency and necessity.

But doesn't Kripke's thought-experiment regarding the discovery of odd atmospheric conditions or corrupting demons also supply a description of a *possible world* in which gold is not yellow? That is, doesn't this thought-experiment, if successful, suffice to secure not only an *epistemic* status for the claim expressed by (1), namely, that it is not known *a priori*, but also a *modal* status for that claim, namely, that it is contingent? The thought-experiment surely at least describes a possible state of the *actual* world, a way that the world *might be* (might yet prove to be), and, we recall, such discoveries would entitle us to conclude both that gold is not yellow, and that gold never was yellow. But does this suffice to demonstrate that we can satisfactorily describe a *possible* world in which gold isn't yellow? One of the most striking features of Kripke's views is that his answer here is "No". Securing an *a posteriori* epistemic status for a given proposition has *no implications at all* regarding the question of its modal status, its necessity or contingency.[2]

The alternative philosophical strategy that lies closest to hand at this point would be to *couple* modality to epistemology and let the necessities fall where they may. Thus one might begin in a pragmatist vein by affirming both that

Whatever is accepted *a posteriori* is defeasible (revisable)

and that

Whatever is defeasible (revisable) is contingent

and conclude straightaway that there are no necessary *a posteriori* truths. I in fact find this sort of epistemically constrained account of modality quite congenial, but rather than claim it straightaway as my own, I shall attribute it, at least temporarily, to a character whom I'll call "the Rival". Kripke's way of treating such modal notions as necessity and contingency, then, is not the Rival's epistemic way. Among the things we need to discover is *why* it isn't and, of course, what it is instead.

Let me therefore turn from Kripke's discussion of gold's color to another of his illustrations, the question of the material composition of a specific table. The sentence

(2) This table is made of wood,

Kripke tells us, an appropriate table having been indicated by suitable pointing gestures, *does* express a necessary truth. We *cannot*, he claims, describe a possible world in which this table—this *very* table, as Kripke frequently puts it—would be and would always have been made, for example, of ice.

In particular, Kripke is convinced that the (material) *substance* of which a material object is composed—as well as its (material) origins—is one of its essential properties.[3] To the extent that he offers us an *argument* for the necessity of (2), it rests initially on successive instantiations of just such a general premise:

(I.0) For any object, x, and for any "kind of stuff" (material), S, if x is made (composed) of S, then x is *necessarily* made of S.

(I.1) For any "kind of stuff", S, if the table is made of S, then the table is *necessarily* made of S.

(I.2) If the table is made of wood, then the table is necessarily made of wood.

(2) The table *is* made of wood.

(I.3) The table is necessarily made of wood.

Given her epistemological inclinations with regard to modal matters, the Rival, we may suppose, is not likely to be moved by this argument. We can rather expect to find her responding somewhat along these lines:

> Of course I grant that *if* (I.0) is true, then so is (I.3). Logic is logic, and the argument given is formally valid as it stands.

But one person's modus ponens is another's modus tollens. I am equally convinced that (I.0) *isn't* true, and, what's more, I can produce an equally valid argument in support of that conclusion. It goes like this: (2) is empirically defeasible; whatever is defeasible is contingent; hence, (I.3) is *false*. Since, however, (2) is true, it follows that (I.2) is false, likewise that (I.1) is false, and hence that (I.0) is false.

One of the especially seductive features of Kripke's new apparatus of "possible worlds" was that it gave the illusion of containing resources for resolving such conflicts between prior convictions regarding modal matters.[4] Thus, confronted with the Rival's disagreement, Kripke's strategy is to transpose the subjunctive question regarding the table in the actual world, "Could this table ("this very table") have been (originally) made of ice?", into an indicative question phrased in terms of "possible worlds": "Can we stipulate or describe or imagine a possible world in which this (very) table *is* (originally) made of ice?" And although our tenacious Rival is inclined to answer this question, too, in the affirmative:

> Here's the stipulation you're looking for:
>
> [S] "Consider a possible world in which this (very) table is (originally) made of ice."

Kripke is now evidently prepared to offer *reasons* for concluding that she has failed to deliver the goods:

> We could conceivably discover that, contrary to what we now think, this table is indeed made of ice from the river. But suppose that it is not. Then, though we can imagine making a table out of another block of wood or even from ice, identical in appearance with this one, and though we could have put it in this very position in the room, it seems to me that this is *not* to imagine *this* table as made of wood or ice, but rather it is to

> imagine another table, *resembling* this one in all external details, made of another block of wood, or even of ice. (113–114)

Now Kripke does acknowledge a contingency in the *neighborhood* of (2). Given a certain "epistemic situation", he says, the table might have *turned out* to have been made of ice, or indeed of anything else. But, contrary to expectations, this truth does not imply the contingency of (2). Rather,

> . . . it means simply that there might have been *a table* looking and feeling just like this one and placed in this very position in the room, which was in fact made of ice. In other words, I (or some conscious being) could have been *qualitatively in the same epistemic situation* that in fact obtains, I could have the same sensory evidence that I in fact have, about *a table* which was made of ice. (142)

According to Kripke, however, that would not be a possible world in which *this* table is made of ice. That would be a possible world in which *another, different,* table—qualitatively quite like this one—is made of ice.

> *This* table itself could not have had an origin different from the one it in fact had, but in a situation qualitatively identical to this one with respect to all the evidence I had in advance, the room could have contained *a table made of ice* in place of this one. (142)

Kripke's analytical apparatus thus evidently requires *inter alia* that we draw a sharp distinction between two species of thought-experiment. On the one hand, there are thought-experiments in which we consider in what ways we might (still, here and now) be mistaken about the actual world. Call these *Type E* (*epistemic*) thought-experiments. In a Type E thought-experiment, we do not, so to speak, venture beyond the actual world. Nothing in such an epistemic thought-experiment involves us in the counterfactual stipulation of a possible world (numerically) *different from* the actual world. On the other hand, there are thought-experi-

ments in which we suppose ourselves to be right about the *actual* world—we stipulatively fix the truths about the actual world to be what we take them to be—and then consider in what ways the world *could have been* different, for example, how its history could have been different from its actual history. Call these *Type M* (*metaphysical*) thought-experiments. A metaphysical thought experiment does involve us in the counterfactual stipulation of possible worlds, worlds *different from* (having a different history from) the actual world.[5]

Kripke's response to the Rival's skepticism regarding his substance essentialism is to argue that she is making one of two characteristic errors:

> *either* she is confusing Type E and Type M thought-experiments,
>
> *or* she is wrong about *which* Type M thought-experiment she is performing, that is, she is actually performing a different Type M thought-experiment from the one she *takes* herself to be performing.

What in either case the Rival has not succeeded in doing, however, is to stipulate or describe or imagine a possible world, (numerically) different from the actual world, in which *this* table is (originally) made not of wood but of ice. Either she has imagined our discovering, in the *actual* world, that this table *never was* made of wood, i.e., that, contrary to what we believe, the table is indeed (and always has been) made of ice from the river—and that is a Type E thought-experiment—or she has imagined a possible world, different from the actual world, in which *some other* table, one made of ice and qualitatively indistinguishable from this one, occupies the (epistemic-evidential) position occupied in the actual world by this table made of wood—and, although that is indeed a Type M thought-experiment, it is the wrong one. The fact that one is able to imagine a possible

world containing some *other* table has no consequences at all regarding the truth or falsehood of modal claims about *this* table.

But have we really made any progress here? For what underwrites Kripke's evident confidence in *these* assertions? The Rival claims, after all, to be imagining a possible world in which *this very table*, not some other one, is originally made of ice. That, after all, is what her stipulation [S] *stipulates*. What warrants Kripke's conviction that, whatever thought-experiments the Rival may successfully consummate, the crucial Type M thought-experiment, one that would establish or confirm the contingency of (2), will not be among them? By Kripke's own lights, after all, possible worlds are stipulated, not discovered by powerful telescopes. Why, then, can't the Rival simply stipulate that, in the possible world that *she* is considering, it is *that very table* that is made of ice, and not some other one?

As far as I can see, Kripke's answer can only be that the Rival isn't imagining the possible world she claims to be imagining because she *can't be* imagining such a possible world. She can't be imagining it, in turn, because there *is* no such possible world; that is, no such world is possible. In other words, Kripke must hold that the Rival's stipulative description of what she *purports* to imagine is contradictory or incoherent. But why is this so? Why, because a thing's (original) material composition is one of its essential properties, of course. What reveals that the stipulation [S] is contradictory or incoherent, in other words, is the fact that it is inconsistent with a *necessary* truth, namely, that the table ("this very table") is (originally) made of wood. It is only because he already accepts (I.0), in short, that Kripke finds the Rival's stipulation [S] inadmissible. But now it is painfully clear that invoking the apparatus of possible worlds has yielded, and indeed can yield, no argumentative advance on the intuitive general essentialist convictions with which Kripke begins.

If Kripke regards (2) as necessary and accepts (I.1) and (I.2), then, this can only be on the grounds of a *prior* and *independent*

judgment to the effect that an object's material composition is one among its essential properties (whereas, for example, a substance's color is not), that is, on the grounds of a prior and independent commitment to the truth of (I.0). The framework of Kripke's "Naming and Necessity" gives us nothing more. Specifically, the necessity of (2) is in no way a *consequence* of Kripke's new analytical apparatus. His only argument for it rests on the assumption (I.0), and (I.0) in turn is not a thesis *for* which Kripke offers any cogent arguments at all. It is a thesis *from* which he argues. It is a prior conviction which is controlling his critique of the "possible world" counterfactual stipulations offered by his skeptical Rival, a critique that could create the appearance, but only that, of giving his necessity thesis independent theoretical support.

Parallel observations pertain to several of Kripke's other claims regarding essential qualities. His ostensible argument for the thesis that Queen Elizabeth's (familial) origins, her parents, could not have differed from those from which she actually came, for example, throws the role of such prior independent judgments into especially sharp relief. Here, too, Kripke begins by conceding the *epistemological* points regarding discoveries which could still, here and now, be made.

> Could she, let's say, have been the daughter instead of Mr. and Mrs. Truman? There would be no contradiction, of course, in an announcement that . . . , fantastic as it may sound, she was indeed the daughter of Mr. and Mrs. Truman. I suppose there might even be no contradiction in the discovery that—it seems very suspicious on either hypothesis that she has a sister called Margaret—that these two Margarets were one and the same person. . . . At any rate, we can imagine discovering all of these things. (112)

Next Kripke stipulates the availability of the relevant empirical antecedent as a true premiss (i.e., assumes that we are right about the actual world): "But let us suppose that such a discovery is not in fact the case. Let's suppose that the Queen really did

come from these parents" (112). Finally, he invokes the apparatus of "possible worlds": "[Can] we imagine a situation in which it would have happened that this very woman came out of Mr. and Mrs. Truman?" His answer, predictably, is "No".

> How could a person originating from different parents, from a totally different sperm and egg, be *this very woman*? One can imagine, *given* the woman, that various things in her life could have changed. . . . One is given, let's say, a previous history of the world up to a certain time, and from that time it diverges considerably from the actual course. This seems to be possible. . . . But what is harder to imagine is her being born of different parents. It seems to me that anything coming from a different origin would not be this object. (113)

However plausible these remarks might strike one as being, their plausibility should not blind us to the absence here of any *argument* for the necessity of origins. The exercise of stipulating possible worlds does not require that one explain *how* a person originating from different parents (i.e., parents other than King George VI and Queen Elizabeth *nee* Lady Elizabeth Bowes-Lyon) could be this very woman (i.e., Queen Elizabeth II), but only that one satisfactorily describe a situation in which that would be the case.

Interestingly enough, the first sentence of this quotation is one of the few places in which the text of the *book Naming and Necessity* diverges from that of the published *essay* "Naming and Necessity". In its original incarnation, the sentence read: "What right would you have to call this baby from completely different parents—in what sense would she be—*this very woman*?"

It is not difficult to understand why Kripke has altered the text at just this point. The unacceptable suggestion of the original wording, of course, was that we here face a problem of "transworld identification", that we must somehow *earn* the right, in describing some possible world, to call the baby "this very woman". But this suggestion runs precisely counter to Kripke's

steady insistence that possible worlds need *not* be described purely qualitatively, but rather may be given by counterfactual stipulations *using* the designators available to us in *this* world:

> [We] do not begin with worlds (which are supposed somehow to be real, and whose qualities, but not whose objects, are perceptible to us), and then ask about criteria of transworld identification: on the contrary, we begin with the objects, which we *have*, and can identify, in the actual world. We can then ask whether certain things might have been true of the objects. (53)

Kripke's supposition that our current beliefs regarding the Queen's parentage are in fact true, in other words, can be put to *use* in his argument only if he has available the requisite additional (tacit) *conditional* premiss to the effect that

> *If* this very woman (i.e., Queen Elizabeth II) came from specific parents (e.g., George VI and the former Elizabeth Bowes-Lyon), *then* it is necessary that she came from those parents.

Now one can produce a Kripkean argument for this requisite premiss, analogous to the earlier argument for (I.2), namely:

(II.0) For any object, x, and for any "origins", O, if x came from O, then x *necessarily* came from O.

(II.1) For any "origins", O, if this very woman (i.e., Queen Elizabeth II) came from O, then it is necessary that she came from O.

(II.2) *If* this very woman (i.e., Queen Elizabeth II) came from specific parents (e.g., George VI and the former Elizabeth Bowes-Lyon), *then* it is necessary that she came from those parents.

Once again, however, the reasoning takes as its point of departure a prior and independent modal conviction, a prior and

independent commitment to the truth of (II.0). What Kripke leaves us with, therefore, is only the bare appeal to the prior judgment with which his ostensible argument regarding Queen Elizabeth concludes: "It seems to me that anything coming from a different origin would not be this object". In short, plausible or implausible, like (I.0), (II.0) is not a thesis for which Kripke argues but a prior conviction from which he argues.

One of Kripke's most exciting and controversial theses is that

> [statements] representing scientific discoveries about what [some] stuff *is* are not contingent truths but necessary truths in the strictest possible sense. It's not just that it's a scientific law, but of course we can imagine a world in which it would fail. Any world in which we imagine a substance which does not have these properties is a world in which we imagine a substance which is not [that stuff], provided these properties form the basis of what that substance is. (125)

Kripke's illustrative case is the elementhood and atomic number of gold. Both

(4) Gold is an element

and

(5) Gold has atomic number 79

he claims, are not merely "physically necessary" truths, but full-fledged "metaphysically necessary" truths, "necessary truths in the strictest possible sense". His ostensible argument, as before, invokes the notion of "possible worlds" in a Type M thought-experiment:

> [Consider] a possible world. Consider a counterfactual situation in which, let us say, fool's gold or iron pyrites was actually found in various mountains in the United States, or in areas of South Africa and the Soviet Union. Suppose that all the areas which actually contain gold now, contained pyrites instead, or

> some other substance which counterfeited the superficial properties of gold but lacked its atomic structure. Would we say, of this counterfactual situation, that in that situation gold would not even have been an element (because pyrites is not an element)? It seems to me that we would not. We would instead describe this as a situation in which a substance, say iron pyrites, which is not gold, would have been found in the very mountains which actually contain gold and would have had the very properties by which we commonly identify gold. But it would not be gold; it would be something else. One should *not* say that it would still be gold in this possible world, though gold would then lack the atomic number 79. It would be some other stuff, some other substance. (124)

Here, for a change, even the Rival will agree with Kripke. But this is not because he has finally produced a cogent argument in support of one of his fundamental substantive essentialist theses. It is rather because he has produced a series of *trivially* necessary truths. Of course, what was found in the mountains in *this* possible world would not be gold, because this possible world is explicitly stipulated to be one in which what was found in the mountains *isn't* gold. We are explicitly enjoined, that is, to "suppose that all the areas which actually contain gold now, contained . . . some *other* substance" (124, my emphasis), and it would be a straightforward contradiction to say *both* that those areas (exclusively) contain gold *and* that they contain (exclusively) some other substance, that is, some substance other than gold.[6]

But does this go any distance toward showing that we cannot stipulate or imagine or describe a possible world in which *gold* isn't an element or doesn't have atomic number 79? Obviously not. The counterfactual stipulations relevant to the necessity or contingency of (4) and (5) would need to begin quite differently, say:

> Consider a possible world in which all the areas which now actually contain *elemental* gold contained *non-elemental* gold instead.

or

> Consider a possible world in which all the areas which now actually contain Au^{79} instead contained a variety of gold having atomic number 82.

Kripke would doubtless object that there can *be* no such "possible worlds", that such counterfactual stipulations are contradictory or incoherent. Not only is the possible world that he has described not a case in which gold might not have been an element, he argues, but (except in the epistemic sense of 'possible') there can be no such case.

> Given that gold *is* this element [with atomic number 79], any other substance, even though it looks like gold and is found in the very places where we in fact find gold, would not be gold. It would be some other substance which was a counterfeit for gold. In any counterfactual situation where the same geographical areas were filled with such a substance, they would not have been filled with gold. They would have been filled with something else. (125)

But, once again, this string of trivial, *analytic* necessary truths:

> Any other substance [i.e., any substance *other than gold*], even though it looks like gold . . . would not be gold,
>
> In any counterfactual situation where the same geographical areas [that are now filled with gold] were filled with such a substance [i.e., a substance *other than gold*], they would not have been filled with gold,

adds no support to Kripke's substantive necessity claims. The trivial fact that we cannot coherently imagine or describe a counterfactual situation in which some element or compound (stipulated to be) *other than gold* is (simultaneously stipulated to be) gold is utterly irrelevant to the question of whether one can stipulate a possible world in which *gold* is (stipulated to be) a

non-elemental substance or a substance with an atomic number other than 79. Once again, the theoretical apparatus of possible worlds, counterfactual stipulations, and Type M thought-experiments drops out of the picture, and we are left with only prior convictions.

We have now confirmed that each of Kripke's specific unconditional essentialist necessity claims depends upon an intuitive *general conditional* essentialist principle. In each case, the principle is instantiated in such a way that its antecedent becomes an empirical, *a posteriori*, truth, and its consequent is then detached by modus ponens as the desired unconditional claim of "metaphysical necessity". If we search for textual or argumentative support for these general conditional essentialist principles themselves, however, we will be disappointed. In each instance, all we will find is that Kripke simply appeals to his prior convictions regarding such matters, to "how it seems to him", or, to invoke a more fashionable phrase, to his "modal intuitions".

It is not obvious that Kripke would find this a *defect* in his presentation, for he evidently takes such intuitions to have considerable probative force: "[Some] philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking." (42). It is not entirely clear how these remarks are properly to be understood. The notion that intuitions function as *evidence* initially suggests that Kripke takes them to play a role in the epistemology of modal theorizing analogous to the role played by *measurements* or *observations* in the case of matter-of-factual, empirical theorizing, but the analogy strikes me as a dubious one. We can, of course, also interpret the notion of "evidence" less literally, and understand Kripke as holding only that modal intuitions function as the best sort of *good reasons* that can be given in support of theoretical or systematic modal principles. This may

well prove to be a defensible view, but it plainly needs some defending.

It remains unclear, for example, why the fact that one finds oneself with an immediate and unreflective inclination to accept a given modal claim is *any sort of reason at all* for believing that claim to be correct. The case of perception provides a useful contrast. Here it is quite possible to explain *why* the fact that *I find myself* with an immediate and unreflective inclination to believe, for example, that there is a cubical red object on a triangular blue table in the center of the room is a *good reason* for me to believe that there is a cubical red object on a triangular blue table in the center of the room. That is, there is a (relatively) straightforward story that can be told regarding how any being possessing my sort of sensory and language-learning capacities can come to be "calibrated", so to speak, as a *reliable indicator* of the colors and shapes of objects in its vicinity. It is far from obvious that the ostensible probative force of modal intuitions allows of being backed in a similar way by analogous explanatory arguments.

Such questions regarding the evidential or probative force of modal (and other) intuitions will gradually gravitate toward the center of our attention. At this stage of my explorations, however, I want to stress only the limited point that assigning even a strong presumptive probative force to intuitions is no help at all when the question at issue is how to adjudicate among *conflicting* intuitions—and this in turn arguably suggests that intuitions as such may, contra Kripke, play only a very minor epistemic role in philosophical theorizing.

In this regard, Hilary Putnam's now-classic "Twin Earth" thought-experiment, originally advanced in support of the conclusion that "meanings aren't in the head", provides an instructive case study.[7] Putnam describes Twin Earth as differing from Earth only in that, wherever we on our planet encounter instances of the natural kind *water*, that is, H_2O, inhabitants of Twin Earth encounter some quite different chemical compound, XYZ.

On Putnam's account, XYZ is *phenomenally* indistinguishable from H_2O, but—although they also *call* it 'water'—the scientifically and technologically more primitive natives of Twin Earth ("Twin Earthlings") lack any analytical chemical and physical techniques adequate to distinguish XYZ *compositionally* from H_2O.

Putnam, in other words, proposes to describe a possible world in which some distinct chemical compound XYZ occupies, so to speak, the natural-kind station occupied on Earth by H_2O, and he concludes, on the basis of this thought-experiment, that "meanings aren't in the head". The argument runs essentially as follows:

(TE1) What's "in the head" of an Earthling when she uses the word 'water' can be identical to what's "in the head" of a Twin Earthling when *she* uses the word 'water'. Suppose it is.

(TE2) H_2O is water, and XYZ isn't water.

(TE3) Since, when an Earthling uses the word 'water' she's referring to H_2O, and hence to *water*, what the Earthling means by 'water' is *water*.

(TE4) But since, when a Twin Earthling uses the word 'water', she's referring to XYZ, and hence *not* to water, what the Twin Earthling means by 'water' *isn't* water.

(TE5) Thus, since the Earthling word 'water' and the Twin Earthling word 'water' differ in meaning, while what's "in the head" of the Earthling is, *ex hypothesi*, identical to what's "in the head" of the Twin Earthling, meanings aren't in the head. QED

The key premiss in this argument is clearly (TE2), and what supports (TE2) are precisely Kripkean intuitions about the necessity attaching to the constitutive or compositional discoveries made by the sciences about such natural substances as gold and water. Kripke himself puts the case this way:

> It certainly represents a discovery that water is H_2O. We identified water originally by its characteristic feel, appearance and perhaps taste, (though the taste may usually have been due to the impurities). If there were a substance, even actually, which had a completely different atomic structure from that of water, but resembled water in these respects, would we say that some water wasn't H_2O? I think not. We would say instead that just as there is a fool's gold there could be a fool's water; a substance which, though having the properties by which we originally identified water, would not in fact be water. And this, I think, applies not only to the actual world but even when we talk about counterfactual situations. If there had been a substance, which was a fool's water, it would then be fool's water and not water. (128)

The last sentence here, of course, is another one of those irrelevant *trivially* necessary truths that we have already encountered in Kripke's discussion of gold. The balance of the passage, however, is precisely an appeal to "intuitions" of the sort that we have found characteristic of Kripke's defense of each of his essentialist theses.

Not only are conflicting intuitions possible here, however, but, viewed from a different angle, Putnam's Twin Earth thought-experiment brings out quite nicely what they are likely to be. For, despite Kripke's contrary conviction ("I think not"), it is clearly possible to regard Putnam as having described a state of affairs in which there would be *two kinds of water*. H_2O would still be water, but XYZ would *also* be water, although a different kind of water. On this view, what it is to be water is, roughly, to be whatever stuff occupies a certain phenomenal natural kind station, and it is always possible that a single such phenomenal natural kind station is, in fact, occupied by more than one (compositional) natural kind structure.[8] 'Water' thus would mean the same thing in English and Twin English, namely *water*, and meanings, although they needn't be, still could be "in the head". Correla-

tively, however, the logical form of the discovery we report by asserting that "water is H_2O" (i.e., of a "theoretical identification"), on this account, would *not* be "F = G" (e.g., "water = H_2O"), which is incompatible with the conjunction of "F = H" (e.g., "water = XYZ") and "G ≠ H", but rather that of a more complicated claim to the effect that the phenomenal properties of water can be explained by positing H_2O in the *role* of water—a claim compatible with the claim that the phenomenal properties of water can *also* be explained by positing XYZ in the role of water.[9] We'll have occasion to return to these considerations later, but it's important to notice, before we proceed, that 'water' is an *everyday* term—not one, so to speak, "owned" by some scientific, linguistic, or philosophical theory. That is why the competing intuitions here initially have, so to speak, epistemic parity. Everyday beliefs, I suggest, are simply noncommittal when it comes to the question of whether there can or cannot be more than one *kind* of water.

The epistemic point is not that this alternative style of account is clearly preferable to Kripke's or, conversely, that his account is clearly preferable to any account of this sort. The point is that, *qua intuition*, the Putnamian-Kripkean conviction that XYZ, if there were such stuff, wouldn't be (another kind of) water, since water is necessarily H_2O, is epistemologically on a par with the contrary conviction that water isn't necessarily H_2O, since XYZ would also be (a different kind of) water. Both convictions are plainly possible,[10] and neither is *obviously* initially more plausible than the other. Each intuition, in turn, harmonizes with a distinct family of more abstract general convictions regarding modal sentences, empirical knowledge, natural kinds, referential discourse, scientific theories, and so on, in their diverse interrelationships. This is the sort of thing that is traditionally called a "philosophical system", and each intuition could, I suppose, be said to "support" the philosophical system into which it naturally "fits". But if such contrary intuitions can have equal probative force, then the probative force of the intuitions *as such*, considered in isolation

from any prior systematic philosophical convictions, must surely be nil. What will be epistemically significant with regard to a philosophical system will not be the fact that it is "supported by intuitions". *Every* such system will initially be equally supported by, so to speak, *its* intuitions. What will be epistemically significant will be the success of the system as a whole in telling a coherent story that diminishes our philosophical puzzlements.

My leading critical idea, if you like, is that Kripke's philosophical system fails on this score. It fails, I shall argue, primarily because it leaves the *epistemology* of its own main concepts and contentions utterly mysterious. Like other views inspired by logico-mathematical achievements—Descartes, Leibniz, Frege, and Russell come to mind—it turns out to be a philosophical system befitting omniscient beings, where epistemological considerations essentially fall by the wayside, but ill suited for understanding the various epistemic accomplishments (*vis-à-vis* our knowledge, for example, of modalities, references, and natural kinds) of the particular sort of *situated* and *perspectival* spatio-temporal creatures that we actually are. Therein, of course, lies a long tale, but our first task is to command a proper survey of Kripke's views. I want consequently to turn next to the concepts and the case that lie at the heart of those views—proper names, rigid designators, and the (putatively) necessary identity of Hesperus and Phosphorus.