

# 8

## Intuition and Modal Error

George Bealer

— ⊕ —

Modal intuitions are not only the primary source of modal knowledge but also the primary source of modal error. An explanation of how modal error arises—and, in particular, how erroneous modal intuitions arise—is an essential part of a comprehensive theory of knowledge and evidence. But, more than that, such an explanation is essential to identifying and eliminating modal errors in our day-to-day philosophical practice. According to the theory of modal error given here, modal intuitions retain their evidential force in spite of their fallibility, and erroneous modal intuitions turn out to be in principle identifiable and eliminable by subjecting intuitions to properly conducted a priori dialectic and theory construction. And, thus, the classical method of intuition-driven philosophical investigation is exonerated.

I begin with a summary of certain preliminaries: the phenomenology of intuitions, their fallibility, the nature of concept-understanding and its relationship to the reliability of intuitions, and so forth. This is followed by an inventory of standard sources of modal error. I then go on to discuss two specific sources: the first has to do with the failure to distinguish between metaphysical possibility and various kinds of epistemic possibility; the second, with the *local* misunderstanding of one's concepts (as opposed to out-and-out misunderstanding, as in Burge's original arthritis case). The first source of error,

— ⊕ —

I am very grateful to Davor Bodrozic, Daniel Demetriou, David Liebesman and, especially, John Bengson and Dan Korman for insightful comments and discussion

190 / **George Bealer**

though much discussed of late, is I believe widely misunderstood; fortunately, it turns out to be comparatively easy to untangle and poses little threat to intuition-driven philosophical investigation. Discussion of the second source, by contrast, has been absent from the philosophical literature. This source of modal error, and the potential to overcome it, has wide-ranging implications for philosophical method. The failure to understand these sources of modal error has recently led to skeptical accounts of intuition and modal error, which are, I will show, ultimately self-defeating.

## 1. Intuition as a Guide to Possibility

First some comments on the nature of intuition. By intuitions we mean *seemings*: for you to have an intuition that  $p$  is just for it to *seem* to you that  $p$ .<sup>1</sup> Here ‘seems’ is understood, not in its use as a cautionary or “hedging” term, but in its use as a term for a genuine kind of conscious episode. For example, when you first consider one of de Morgan’s laws, often it neither seems true nor seems false; after a moment’s reflection, however, something happens: it now just seems true. This kind of seeming is *intellectual*, not experiential—sensory, introspective, imaginative.

Intuition is different from belief: you can believe things that you do not intuit (e.g., that Paris is in France), and you can intuit things that you do not believe (e.g., the axioms of naive set theory). The experiential parallel is that you can believe things that do not appear (seem sensorily) to be so, and things can seem sensorily in ways you do not believe them to be (as with the Müller-Lyer arrows). Moreover, intuition is typically prior to belief in the order of discovery and evidence: until Putnam we did not even have beliefs about twin earth, but directly upon encountering the example most of us had the intuition that there would be no water on twin earth and only thereafter formed the associated belief. Now, since intuition is analogously different from other psychological attitudes (judging, guessing, imagining, etc.) and from common sense, I believe there is no choice but to accept that intuition is a *sui generis* propositional attitude.

<sup>1</sup> I owe this point to George Myro. Here and in certain other places I use propositional variables, e.g., ‘ $p$ ’ where sentential variables, e.g., ‘ $A$ ’ are strictly speaking called for. Confusion should not result.

**Intuition and Modal Error / 191**

The sort of intuitions relevant to the a priori disciplines are *rational* intuitions (intuitions that present themselves as necessary), not *physical* intuitions. According to traditional usage, “thought experiments” appeal, not to rational intuitions, but to physical intuitions (and the like). Here one constructs hypothetical cases about which one tries to elicit, say, intuitions deriving from one’s implicit mastery of relevant physical laws—as, for example, in Newton’s bucket thought experiment: is it physically possible for the fluid to remain perfectly flat? Not according to physical intuition. Is it metaphysically possible? Of course. Unlike physical intuition, rational intuition derives from one’s understanding of one’s concepts, not of empirical laws. Does this imply that a priori knowledge is always the result of *conceptual analysis*? No, not unless the latter includes various necessities that traditionally were thought to be synthetic, not analytic.

The set-theoretic paradoxes establish an important moral, namely, that intuition can be fallible and a priori belief, revisable—contrary to early modern epistemological dogma. We must therefore embrace the alternative tradition—reaching from Plato to Gödel—that recognizes that a priori justification is fallible and holistic, relying respectively on dialectic and theory construction.

It is our standard epistemic practice to count intuitions as evidence, or reasons, absent special reason not to do so (much as we count ostensible sense perceptions as evidence absent special reason not to do so). But (I argue) we have no such special reason in the case of intuitions. Moreover, if we denied, without any special reason, that intuitions are evidence, we would land in an epistemically self-defeating situation, and, therefore, there is no rational alternative but to accept intuitions as evidence. Finally, since modal intuitions are not, in respects relevant to evidential status, different from nonmodal intuitions, they in particular must be accepted as evidence. (I have defended each of these points elsewhere and will just assume them here.<sup>2</sup> I will, however, provide a discussion in §8 of the self-defeating nature of denying the evidential status of intuition as it arises in the context of modal error and scientific essentialism.)

These conclusions raise two questions. First, *why* are intuitions evidence (reasons)? The answer lies in a reliabilist theory of basic sources of evidence.

<sup>2</sup> For detailed arguments, see George Bealer, “The Incoherence of Empiricism”, *Aristotelian Society*, supplementary volume, 66 (1992), 99–138, and other papers mentioned below.

192 / **George Bealer**

(Reliabilist theories of nonbasic sources of evidence, on the other hand, face standard counterexamples.) Contingent reliabilist theories of basic sources of evidence, according to which there is a contingent nomological tie between the deliverances of such sources and the truth, are open to counterexamples: for instance, contrary to contingent reliabilism, telepathically generated guessing would not count as a basic source of evidence even if it had a nomologically reliable tie to the truth. The explanation, rather, is provided by *modal reliabilism*—the doctrine that there is a certain kind of qualified modal tie between basic sources of evidence and the truth. Intuition (like phenomenal experience) meets this condition.

Second, why should there be a qualified modal tie between intuitions and the truth? The explanation is provided by an analysis of what it is to understand one's concepts. The intuitive idea underlying the analysis is that the identity of one's concepts is manifested in the intuitions involving those concepts that one would have as one approaches cognitively ideal conditions (intelligence, attentiveness, memory, etc.). According to the resulting explanation, the qualified modal tie between intuitions and the truth does not have a supernatural source (as perhaps it does in Gödel's theory of mathematical intuition); rather, it is simply a consequence of what it is to understand the concepts involved in one's intuitions.

I have defended these positions in other work.<sup>3</sup> Here I will be concerned to explain how erroneous modal intuitions arise. The indicated analysis of understanding concepts will play a critical role in my account of modal error. I will return to it after dealing with three further preliminaries.

## 2. Concrete-Case Intuition, Conceivability, and Metaphysical vs. Epistemic Possibility

### 2.1 Concrete-Case Intuition

Theoretical intuitions often play a central role in philosophical debates. Consider, for instance, the theoretical intuitions many people have that free will and determinism are incompatible or that causal closure entails

<sup>3</sup> See, e.g., George Bealer, "Philosophical Limits of Scientific Essentialism," *Philosophical Perspectives* 1 (1987), 289–365; "A Priori Knowledge and the Scope of Philosophy", *Philosophical Studies* 81 (1996), 121–42; and "A Theory of the A Priori", *Philosophical Perspectives* 13 (1999), 29–55.

## Intuition and Modal Error / 193

epiphenomenalism. Theoretical intuitions, however, are typically far more fallible than concrete-case intuitions—much as observations that are heavily theory-laden are much more fallible than those that are not. (This is why in controversies over physical measurement, disputes are typically adjudicated by comparatively non-theory-laden observations such as where the arrow is pointing on the dial or where the endpoints of a rod are on the yardstick.) In contemporary philosophy, it seems that some people are content to found their philosophical theories on a few central theoretical intuitions rather than concrete-case intuitions. This practice lacks historical perspective. The history of philosophy is littered with examples of philosophers no less brilliant than our contemporaries who founded their philosophy on small families of theoretical intuitions which they found especially compelling. The list is embarrassingly long—ranging from Parmenides to Berkeley and Hume, Spinoza and Leibniz, Hegel and Bradley, and on to Schlick and Ayer. To one group of philosophers, certain theoretical principles can seem self-evident whereas to an opposing group the opposites can seem just as compelling. Left at this level, a “battlefield of endless controversies,” as Kant puts it, is inevitable.

The only solution is to defer to concrete-case intuitions. (This is so even in the case of widely held general principles such as the transitivity of the part/whole relation and perhaps the law of excluded middle or even Leibniz’s Law.) This is not to say that these intuitions will themselves be in harmony; the point is that significant overlapping collections of them are in sufficient harmony to adjudicate the dispute. We witness success at this in an impressive list of cases that we now all take for granted. For example, the perceptual-relativity refutation of phenomenalism; the Spartan-pretender refutation of logical behaviorism; the defective-instrument refutation of instrumentalism; the unrepeatable-events refutation of the inductivist theory of justification; and on and on. It is, however, only hubris to think that at just this point in history our theoretical intuitions are at last reliable.

### 2.2 Conceivability

Many philosophers take conceivability and inconceivability to be the primary guide to possibility.<sup>4</sup> I think this is a mistake—well, unless when you say

<sup>4</sup> For a more detailed discussion of conceivability, see George Bealer, “Modal Epistemology and the Rationalist Renaissance”, in Tamar Szabo Gendler and John Hawthorne (eds.), *Conceivability and Possibility* (New York: Oxford University Press, 2002), sect. 1.2.

194 / **George Bealer**

‘It is conceivable that p’, all you are saying (at least conversationally) is that you have an intuition that p is possible; and when you say ‘It is inconceivable that p’, all you are saying is that you have an intuition that it is impossible that p. If so, a lot of confusion would be avoided if we simply talked about possibility and impossibility intuitions. The same goes for ‘imaginable’ and ‘unimaginable’.

Suppose, however, that this easy idiomatic gloss on ‘conceivable’ and ‘inconceivable’ is not correct and that these terms are instead taken at face value as literal expressions of certain modal facts: it is conceivable that p iff it is possible for someone to conceive that p; it is inconceivable that p iff it is not possible for someone to conceive that p. Then we have a pair of problems. First, unlike intuitions of possibility and impossibility, conceivability and inconceivability would not be suited to play their reputed evidential role in modal epistemology. That it is possible, or impossible, to conceive that p is itself a modal fact. But in order for someone to acquire evidence (reasons), something must *actually happen*: a datable psychological episode must *occur* (the occurrence of a sensation, an introspective or imaginative experience, a seeming memory, an intuition). Modal facts do *not* occur. Nothing *happens* when something is conceivable or inconceivable. So something’s merely being conceivable or inconceivable cannot provide anyone with evidence (reasons) for anything.

Second, our beliefs about what is conceivable, or inconceivable, can be highly inferential and are often theoretical. True, one way you can come to believe that it is possible for someone to conceive that p is for *you actually* to conceive that p. But why should your conceiving that p provide you with evidence that p is possible? I can see no reason why it should unless conceiving that p involves intuiting that p is possible. (Stephen Yablo himself says, “In slogan form: *conceiving involves the appearance of possibility*.”<sup>5</sup> But then one is right back to relying on modal intuitions. The moral is simple: talk of conceivability and inconceivability invites (avoidable) confusion. The same goes for imaginability and unimaginability.

<sup>5</sup> Stephen Yablo, “Is Conceivability a Guide to Possibility?”, *Philosophy and Phenomenological Research*, 53 (1993), 5. In support of the centrality of modal intuition, Yablo tells us that “modal intuition *must* be accounted reliable if we are to credit ourselves with modal knowledge . . .” (“The Real Distinction between Mind and Body”, *Canadian Journal of Philosophy*, supp. vol. 16 (1990), 179).

### 2.3 Metaphysical Possibility and Epistemic Possibility

The modal expressions ‘could’, ‘can’, ‘might’, and ‘possible’ are used in diverse ways which fall into two broad classes: epistemic and nonepistemic. (An analogous division holds for ‘must’ and ‘necessary’.) In modal logic, metaphysics, and philosophy of language and mind, the primary focus is on a certain form of nonepistemic necessity—in Kripke’s words, necessity *tout court*. Kripke christened this necessity ‘metaphysical necessity’.

To illustrate some of the epistemic uses of ‘could’, consider any thinkable necessary truth  $p$ .<sup>6</sup> The first use is the ‘could’-of-ignorance: absent what we deem to be adequate evidence (or adequate justification) one way or the other about  $p$ , we can truly say, “It could be that  $p$ , and it could be that not  $p$ . We just do not know yet.” (For example, this can be truly said of Goldbach’s Conjecture.) But once we have adequate evidence (justification) one way or the other, what was meant in speaking *that way* can no longer be truly said. Second, there is the ‘could’-of-less-than-complete-certainty: if we have less than complete certainty about  $p$  (even if we have adequate evidence, or justification, for  $p$ ), we can still truly say, “We still could be mistaken; we know we can be wrong about almost anything.” (For example, even though we now have a proof of Fermat’s Last Theorem, this can still be truly said of it.) Third, there is the ‘could’-of-qualitative-evidential-neutrality (as I will call it): for a posteriori necessities, we can often truly say, “It could have turned out that  $p$ , and it could have turned out that not  $p$ .” And this is so, even though, meant this way, this cannot be said of any traditional a priori necessities. For example, meant this way, ‘Whether Hesperus was Phosphorus could have turned out either way’ would be true, even though when meant the same way ‘Fermat’s Last Theorem could have turned out either way’ would be false. (Of course, there is a corresponding use of ‘It could turn out either way’.)

<sup>6</sup> Importantly, these uses of ‘could’ need not correspond to distinct *literal meanings*; it is enough that they be standard uses of the term in the sort of ordinary contexts relevant to modal epistemology. The ensuing discussion of these epistemic uses of ‘could’, while inspired by Kripke’s comments in *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1980, 103–5, 140–4), departs from Kripke insofar as he evidently did not regard these uses as established standard uses of ‘could’ in ordinary language. For a more detailed discussion of these uses, as well as a critique of various alleged uses of ‘could’ (e.g., the alleged ‘could’-of-‘logical’-possibility), see my “Modal Epistemology and the Rationalist Renaissance” (77–79).

196 / **George Bealer**

A few semi-formal remarks about these epistemic uses of ‘could’ might be helpful. Suppose someone intends the ‘could’-of-ignorance when uttering the sentence ‘It could be that  $p$ ’ in some relevant conversational context.<sup>7</sup> Then, the asserted proposition would be the proposition that results when an associated propositional operation  $\diamond_{\text{ignorance}}$  is applied to the proposition  $p$ . (In symbols:  $\diamond_{\text{ignorance}} p$ .) The truth conditions of the resulting proposition would be as follows: the proposition that  $\diamond_{\text{ignorance}} p$  is true iff it is unknown whether  $p$ .

The ‘could’-of-less-than-complete-certainty may be represented with the operator ‘ $\diamond_{\text{uncertainty}}$ ’. The truth conditions would be: the proposition that  $\diamond_{\text{uncertainty}} p$  is true iff it is not completely certain that  $p$ .

Finally, the ‘could’-of-qualitative-evidential-neutrality may be represented with ‘ $\diamond_{\text{q-e-n}}$ ’. The truth conditions (inspired by Kripke’s comments in *Naming and Necessity*) would be: the proposition that  $\diamond_{\text{q-e-n}} p$  is true iff it is possible for there to be a population  $c$  with attitudes toward  $p$  and it is possible for there to be a population  $c'$  whose epistemic situation is qualitatively identical to that of  $c$  such that the proposition  $p'$ , which in  $c'$  is the epistemic counterpart of  $p$  in  $c$ , is true.<sup>8</sup>

(Contrast the foregoing with the untenable account of epistemic possibility offered by David Chalmers.<sup>9</sup>)

Note that in each of these three biconditionals the proposition denoted by the left-hand side (e.g., that  $\diamond_{\text{q-e-n}} p$ ) need not be *identical* to that which is expressed by the associated right-hand side; indeed, in many cases they are intuitively different. This feature allows the above account to avoid various difficulties that undermine other accounts of epistemic uses of ‘could’. A case in point is Kripke’s account of the ‘could’-of-qualitative-evidential-neutrality,

<sup>7</sup> Here and certain other places I use single quotation marks where, strictly, corner quotation marks are required.

<sup>8</sup> In symbols:  $(\exists c)[\text{Attitudes}(c, p) \ \& \ (\exists c')(\exists p')[\text{QualitativelyIdentical}(c', c) \ \& \ \text{Counterparts}(\langle p', c' \rangle, \langle p, c \rangle) \ \& \ \text{True}(p')]]$ .

<sup>9</sup> David Chalmers, “Does Conceivability Entail Possibility?”, in Szabo Gendler and Hawthorne (eds.), *Conceivability and Possibility*, 145–200. For instance, on Chalmers’s account ‘It could turn out that  $A$ ’ is true iff it is not metaphysically possible for an ideally rational being to know a priori that not  $A$  (see, e.g., pp. 157 and 162). Likewise for ‘It is epistemically possible that  $A$ ’. But such accounts are mistaken. Is it *epistemically* possible—would you say that it could turn out epistemically—that there are no beliefs? Certainly not. But it is not metaphysically possible for any being to have a priori knowledge that there are beliefs. So, Chalmers’s account wrongly implies that it is epistemically possible—it could turn out epistemically—that there are no beliefs.

Intuition and Modal Error / 197

for as we will see, the above account sidesteps various problems confronting Kripke's account.

### 3. Understanding Concepts

Let us return to the prospect (mentioned in §1) of an analysis of what it is to understand a concept. There is a weak nominal sense in which a person can be said to possess a concept:

A subject possesses a given concept at least nominally iff the subject has natural propositional attitudes toward propositions which have that concept as a constituent content.

Possessing a concept in this sense is compatible with what Tyler Burge calls *misunderstanding* and *incomplete understanding* of a concept ('misunderstanding' for cases where there are errors in the subject's understanding of the concept and 'incomplete understanding' for cases where there are gaps).<sup>10</sup> Possessing a concept in the nominal sense is also compatible with, for example, having propositional attitudes merely by virtue of the attribution practices of third-party interpreters. But possessing a concept in such ways, or modes (merely nominally, incompletely, erroneously, and so forth), is very different from *genuinely understanding* the concept. The goal is to analyze this notion.

The intuitive idea behind the intended analysis of understanding is that the identity of one's concepts is manifested in the intuitions involving those concepts that one would have as one approaches cognitively ideal conditions. Specifically, for target concepts  $c$  and relevant hypothetical test cases  $p$  involving  $c$ ,<sup>11</sup> if the subject understands  $c$  and the auxiliary concepts involved in  $p$ , as the subject approaches ideal cognitive conditions (intelligence, attentiveness, memory, etc.), the subject's intuitions regarding the applicability of  $c$  to  $p$  become increasingly *truth-tracking*. That is, as the subject approaches ideal cognitive conditions, it would be increasingly the case that, for such  $c$  and  $p$ , the subject would have the intuition that  $c$  applies to  $p$  if and only if  $c$  really does apply to  $p$ . If the subject did not have such truth-tracking

<sup>10</sup> Tyler Burge, "Individualism and the Mental", *Midwest Studies in Philosophy*, 4 (1979), 73–122.

<sup>11</sup> A hypothetical case  $p$  involving  $c$  is a proposition of the form: it is possible that, for some  $x$ ,  $x$  is in such and such concrete hypothetical situation and  $x$  falls under the concept  $c$  in that situation.

198 / George Bealer

intuitions, the right thing to say would be that either the subject does not genuinely understand one or more of the concepts involved, or the subject's cognitive conditions are not really those indicated, or some other such (perhaps presently unforeseeable) defeater is present.<sup>12</sup>

On this picture, when a subject's mode of understanding shifts to genuine understanding from, say, an incomplete and incorrect understanding, there is an associated shift in the subject's intuitions—in both *quantity* and *quality*. The quantity grows because one's understanding is no longer incomplete, that is, the gaps in one's understanding responsible for the “don't knows” are filled. The quality improves because one's understanding is no longer erroneous, that is, errors in one's understanding are corrected.

Intuition thus bears a qualified modal tie to the truth. The indicated tie does not have a supernatural source; rather, it is simply a consequence of what it is to understand the concepts involved in one's intuitions. Genuine understanding of one's concepts consists in possessing them in a mode such that at the end of a certain idealized intuition-driven process (conducted in ideal cognitive conditions and during which the chosen test propositions are all genuinely understood) converges on relevant necessary truths in a certain fashion. The analysis of understanding one's concepts thus implies that (and, hence, explains why) there should be such a tie to the truth at the end of a priori theorizing.

A common criticism of using the analysis of understanding to explain intuition's tie to the truth is that it amounts to invoking a “dormitive virtue,” which is either unacceptably mysterious or viciously circular.<sup>13</sup> The objection fails, however, for in the present context the explanandum is a modal fact—i.e., intuition's qualified *necessary* tie to the truth. And necessities call for a very different sort of explanation from that called for by contingencies. In the explanation of necessities, it is wholly appropriate to articulate essences, and it is of the essence of the understanding of concepts that intuitions involving those concepts be correct (given ideal cognitive conditions, notably intelligence). This is compatible with its being of the essence of intelligence to have the complementary property. In fact, this complementarity is paradigmatic of functionally

<sup>12</sup> I should emphasize that this is only a rough sketch of the final analysis; it omits crucial details and the supporting justification. For a detailed discussion and defense, see “A Theory of the A Priori”.

<sup>13</sup> See, e.g., Paul Boghossian, “Knowledge of Logic”, in Paul Boghossian and Christopher Peacocke, *New Essays on the A Priori* (Oxford: Clarendon Press), 229–54.

**Intuition and Modal Error / 199**

definable families of basic properties. Thus, if the contemplated criticism should work against the proposed explanation of intuition's tie to the truth, it should by parity work against these other functional definitions (and implicit-turned-direct definitions and perhaps impredicative definitions generally). But plainly this is not so. Evidently, then, it is safe to make use of the analysis of understanding in order to explain intuition's qualified modal tie to the truth.

#### 4. Modal Error

I have emphasized that intuition is fallible—but not so fallible as to undermine it as a source of evidence. To assure ourselves that this is so, it is appropriate to inventory the sorts of error to which modal intuition succumbs and to explain what is going wrong. This, in turn, would help us to identify what conditions need to be optimized in order to eliminate, or least confine, modal intuitional errors.

Many modal intuitional errors have the same etiology as nonmodal intuitional errors. Consider, for example, the following five sources:

*Cognitive deficiency.* Some errors have their origin in cognitive deficiencies (intelligence, attentiveness, etc.). For example, such deficiencies explain why many people initially have the erroneous modal intuition regarding the Barber Paradox (i.e., that it is possible for someone to shave all and only those who do not shave themselves).

*Conflation.* Some errors are born from a failure to take note of various conceptual or logical distinctions. We see this, for example, in Galileo's paradox of infinity (which is based on the common intuition that there cannot be as many odd numbers as there are natural numbers); here two concepts of equinumerosity are conflated, namely, the concept of one-to-one correspondence and the concept of proper inclusion.

*Misunderstanding.* Some errors have to do with out-and-out misunderstanding of one's concepts—as, for example, in the case of the positive intuition Burge's arthritis man would have if asked whether it is possible to have arthritis in the thigh.

*Underdescription and context-sensitivity.* Some errors arise from underdescription of the case under consideration or from inattention to relevant contextual factors (including Gricean pragmatic factors and context-dependent norms involved in the specification of the case).

200 / **George Bealer**

*Theoreticality.* As discussed in §2, theoretical intuitions (as opposed to concrete-case intuitions) are typically more vulnerable to error, much as theory-laden observations are.

In what follows, I want to focus on two particularly thorny sources of modal intuitional error: the first has to do with the failure to distinguish between metaphysical possibility and various kinds of epistemic possibility; the second, with the *local* misunderstanding of one's concepts (as opposed to out-and-out misunderstanding, as in Burge's original arthritis case). Untangling these two types of modal error plays a pivotal role in the defense of scientific essentialism (hereafter SE)—for example, in the defense of the thesis that water, gold, and so forth are natural kinds having microscopic (or otherwise hidden) a posteriori essences. Our discussion of the first source of modal error (§§5–6) is wholly concerned with people, like Kripke himself, who have pro-SE modal intuitions but who also report having apparently anti-SE modal intuitions. Our discussion of the second source of modal error (§7) is instead concerned with people who have anti-SE intuitions but who simply lack pro-SE intuitions.

## 5. Modal Error and Rephrasal Strategies

All successful arguments for SE rely on intuitions, for example, intuitions concerning Aristotle and the teacher of Alexander, water and the water-like stuff on twin earth, and so forth. Indeed, careful reflection on the structure of such arguments (discussed in §6.2 below) reveals that without the evidential support of intuitions SE would be unjustified.<sup>14</sup> But there is a *prima facie* problem with the reliance on intuition, for Kripke and those who share his pro-SE intuitions also have a host of apparently anti-SE intuitions, for example, the intuition that it could have turned out that some samples of water contained no hydrogen. Kripke, of course, was well-aware that he needed to deal with this problem in order for his defense of SE to succeed. He did so by developing an account of such apparent errors in modal intuition.

<sup>14</sup> To be sure, certain contemporary advocates of SE might wish to abandon the intuitional defense of SE altogether. But this results in an essentially unstable position: absent intuition, radical empiricist Quineanism about modality, not scientific essentialism, is the result. (This difficulty will play a role in §8.) In the present context, we will only be concerned with philosophers who do not wish to abandon intuition-driven philosophical method.

**Intuition and Modal Error / 201**

Our discussion of such modal errors will proceed by presenting, and then assessing, Kripke's account.

So what are Kripke and those who share his pro-SE intuitions to make of the apparent conflict? Bear in mind that this group includes not just proponents of SE but also critics. Proponents of SE have two responses.

First, they could simply declare that anti-SE intuitions are mistaken whereas their own pro-SE intuitions are correct. But critics of SE could simply meet this response by claiming that things are the other way around. The result would be a stalemate.

The second response is, of course, to try to resolve the apparent conflict. The leading general strategy for doing this follows Kripke in deeming the widespread conflict among our intuitions to be only an appearance. All, or most, of our intuitions are correct. (Indeed, Kripke tells us, "I think [intuition] 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."<sup>15</sup> Kripke also seems to believe that our intuitions must be on the whole correct if scientific essentialism is to be based on adequate evidence.) Despite their correctness, however, many are *misreported*. When we rephrase our (apparently) anti-SE intuitions to make them consistent with our pro-SE intuitions, we succeed. But, of course, this is not enough. Two further requirements (not discussed by Kripke) must be met. First, when opponents of SE try to rephrase the pro-SE intuitions to make them consistent with the apparently anti-SE intuitions, they fail, thus provisionally breaking the impending stalemate in favor of SE (see below for an illustration). Second, it must be the case that there is not some further, equally plausible, rephrasal strategy that is asymmetric in this way, but this time favoring anti-SE rather than pro-SE. For, if there were such a rephrasal strategy, the impending stalemate would be restored.

Kripke and his followers have used two rather different rephrasal strategies. The first turns on an alleged equivocation involving a confusion about features of our epistemic situation. According to this strategy, when we report our pro-SE intuitions (e.g., twin-earth intuitions), what we say is strictly and literally true; but when we report our apparently anti-SE intuitions, we confuse ordinary possibility with the possibility of a certain kind of epistemic situation. For example, when we say 'It could have turned out that some

<sup>15</sup> *Naming and Necessity*, 42.

202 / **George Bealer**

samples of water contained no  $H_2O$ ', what we say is strictly and literally false. The intuition is true but incorrectly reported. Kripke develops this idea in connection with the Hesperus/Phosphorus example:

Now this seems very strange because in advance, we are inclined to say, the answer to the question whether Hesperus is Phosphorus might have turned out either way. And so it's true that given the evidence that someone has antecedent to his empirical investigation, he can be placed in a sense in exactly the same situation, that is a qualitatively identical epistemic situation [to ours], and call two heavenly bodies 'Hesperus' and 'Phosphorus', without their being identical. So in that sense we can say that it might have turned out either way.<sup>16</sup>

Generalizing from these examples, we arrive at the following schema. The true thing incorrectly reported by 'It could have turned out that A' is correctly reported by the following sentence: It is possible that a population of speakers in an epistemic situation qualitatively identical to ours would make a true statement by asserting 'A' with normal literal intent. Consider the true intuition that we incorrectly report with 'It could have turned out that there were samples of water containing no  $H_2O$ '. The rephrasal comes out true because in the envisaged population of speakers 'water' might not name water but rather XYZ (or 'H' might not name hydrogen but perhaps X instead). When rephrased thus, the original apparently anti-SE intuition is plainly consistent with the thesis that, necessarily,  $water = H_2O$ .

Kripke's second rephrasal strategy is this.<sup>17</sup> Suppose that ' $R_1$ ' and ' $R_2$ ' are co-designating rigid designators whose designatum was fixed by the nonrigid (i.e., contingent) designators ' $D_1$ ' and ' $D_2$ ', respectively. When we report an apparently anti-SE intuition with 'It could have turned out that  $R_1 \neq R_2$ ', our intuition is correct but misreported. It is correctly reported with 'It is possible that  $D_1 \neq D_2$ '. On its standard narrow-scope reading, the latter sentence is consistent with the SE thesis that, necessarily,  $R_1 = R_2$ . For ' $D_1$ ' and ' $D_2$ ' are only contingently co-designating. For example, on this proposal 'It could have turned out that  $water \neq H_2O$ ' might be rephrased as: 'It is possible that the clear thirst-quenching stuff  $\neq$  the such-and-such chemical compound'. The latter is consistent with the thesis that, necessarily,  $water = H_2O$ , for there is

<sup>16</sup> *Naming and Necessity*, 103–4.

<sup>17</sup> It is the second rephrasal strategy, not the first, that plays a pivotal role in Kripke's argument (in *Naming and Necessity*) against the identity theory in philosophy of mind. This argument is critically assessed in George Bealer, "Mental Properties", *Journal of Philosophy*, 91 (1994), 185–208.

Intuition and Modal Error / 203

a possible situation in which there is a unique clear thirst-quenching stuff that is not a such-and-such chemical compound.

Both rephrasal strategies are flawed. First, let us consider two problems with the second rephrasal strategy. (I will criticize the first strategy in §6.1.)

(1) It is based on the thesis that, when we report an intuition with ‘It could have turned out that  $R_1 \neq R_2$ ’, often the true thing we have in mind is strictly and literally reported with ‘Possibly,  $D_1 \neq D_2$ ’, where ‘ $R_1$ ’ and ‘ $R_2$ ’ are names and ‘ $D_1$ ’ and ‘ $D_2$ ’ are descriptions. But Kripke, of all people, should not be proposing that, when we make use of a proper-name sentence in ordinary conversation (even if the sentence happens to be of the form ‘It could have turned out that  $R_1 \neq R_2$ ’), we have in mind something *descriptive*. After all, the situation is *phenomenologically and behaviorally indistinguishable* from situations in which we have in mind something *nondescriptive* (as, for example, when Kripke asserts his well-known thesis ‘If Hesperus = Phosphorus, then it is not possible that Hesperus  $\neq$  Phosphorus’). For Kripke to deny this would be *ad hoc* and implausible. Hence, the rephrasal strategy itself is implausible.

(2) It can be shown that this rephrasal strategy does not even accomplish the goal of breaking the impending stalemate between our apparently conflicting pro- and anti-SE intuitions. Specifically, this rephrasal strategy lacks the requisite asymmetry property (described above). For one can wield it so as to sustain the original force of prima facie anti-SE intuitions and to deflate the original force of the pro-SE intuition reports, thereby rendering our prima facie pro-SE intuitions consistent with the rejection of SE. The following recipe provides one way of doing this. Adopt the traditional description theory of names. Hold that names occurring in reports of anti-SE intuitions *are* being used strictly and literally and that they express *nonrigid* descriptive content. Hold that names occurring in reports of pro-SE intuitions are *not* being used strictly and literally and that they are being used to express *rigid* descriptive content. (E.g., this rigidity could be the result of implicitly understood actuality-operators.) The rephrasal strategy can thus be used to affirm anti-SE just as effectively as it can be used to affirm pro-SE. Hence, the impending stalemate is not broken.

One advantage of the first rephrasal is that, unlike the second, it does have the requisite asymmetry. Because our anti-scientific-essentialists are Lockean internalists, they are committed to holding that the meaning of ‘water’ and other relevant expressions cannot differ across populations of speakers in qualitatively identical epistemic situations. Accordingly, they must hold that their rephrasal of the pro-SE intuition report *entails* the pro-SE report itself.

204 / **George Bealer**

(Consider the apparently pro-SE intuition reported with ‘Possibly, there is a twin earth such that . . . the clear thirst-quenching samples are not samples of water’. When the first rephrasal strategy is applied, such anti-scientific-essentialists must hold that this intuition is true but incorrectly reported; it is correctly reported with ‘It is possible for there to be a population of speakers in an epistemic situation qualitatively identical to ours who would make a true statement by asserting “There is a twin earth such that . . . the clear thirst-quenching samples are not samples of water” with normal literal intent’. But, given their Lockean internalism, these anti-scientific-essentialists must hold that such a population of speakers would mean what we mean with ‘The clear thirst-quenching samples are not samples of water’. If so, this rephrasal entails that it is metaphysically possible that there be a twin earth such that . . . the clear thirst-quenching samples are not samples of water.<sup>18</sup> Consequently, our anti-scientific-essentialists are committed to holding that the rephrasal has the same pro-SE force as the original report.) By contrast, scientific essentialists are not traditional internalists, so they are free to hold that the meaning of ‘water’ and other relevant expressions can differ across populations of speakers in qualitatively identical epistemic situations. So when the original intuition seems to have an anti-SE force, they are free to hold that that force is deflated upon rephrasal. The impending stalemate is thus broken in their favor.

No known competing rephrasal strategies have the requisite asymmetry.<sup>19</sup> I have space to show why for just one of these strategies, though the other cases are very similar. This strategy, which is advocated by Thomas Nagel, Michael Levin, and Crispin Wright, involves the idea of imaginative projection, as I will call it.<sup>20</sup> The idea is that when we assert ‘It could have turned out that

<sup>18</sup> After all, the rephrasal is equivalent to ‘It is possible for there to be a population of speakers in an epistemic situation qualitatively identical to ours in whose language there is a *true* sentence synonymous to the English sentence “There is a twin earth such that . . . the clear thirst-quenching samples are not samples of water”’, and this sentence plainly entails the original report ‘There is a twin earth such that . . . the clear thirst-quenching samples are not samples of water’.

<sup>19</sup> Chalmers’s approach would have the requisite asymmetry but, as we saw (in note 9), is unsatisfactory on other grounds.

<sup>20</sup> Thomas Nagel (“What Is It Like to be a Bat?”, *Philosophical Review*, 83 (1974), 435–50). Michael Levin (“Tortuous Dualism”, *Journal of Philosophy*, 92 (1995), 314–23). Crispin Wright, (“The Conceivability of Naturalism”, in Szabo Gendler and Hawthorne (eds.), *Conceivability and Possibility*, 437–8). Christopher Hill critically assesses the imaginative projection strategy in “Imaginability, Conceivability, Possibility and the Mind-Body Problem” (*Philosophical Studies*, 87 (1997), 65–72).

**Intuition and Modal Error / 205**

water  $\neq$  H<sub>2</sub>O', the true thing we have in mind is: it is possible to imagine what it would be like to uncover evidence that would show that water  $\neq$  H<sub>2</sub>O. But this rephrasal strategy fails because it lacks the requisite asymmetry. Specifically, opponents of SE can use it to deflate their prima facie pro-SE intuitions just as effectively as advocates of SE can use it to deflate their prima facie anti-SE intuitions. Consider, for example, the pro-SE twin-earth intuition reported by 'It could have turned out that there is a twin earth macroscopically like earth but where the water-like samples would not be water'. By applying the present rephrasal strategy, opponents of SE would arrive at the following: it is possible to imagine what it would be like to acquire evidence showing that a certain macroscopic duplicate of water (e.g., XYZ) is not water. (For example, I can imagine what it would be like for such evidence to emerge in the course of a partly scientific and partly philosophical investigation.) Since the imaginability of such a scenario does not entail that it is metaphysically possible for there to be a macroscopic duplicate of water that is not water, this rephrasal deflates the original pro-SE intuition. So stalemate would ensue. Ironically, this rephrasal strategy turns out to be as much a threat to SE as it is an aid.<sup>21</sup> (The resulting dialectical situation thus resembles that which confronts advocates of skeptical accounts of modal error, to be discussed in §8.)

## 6. Modal Error and Epistemic Possibility

### 6.1 No Conflict

I used to think that the upshot of the discussion in §5 was that Kripke's first rephrasal strategy (and no other) successfully deflates our prima facie anti-SE intuitions and thus reconciles the apparent conflict.<sup>22</sup> But I have come to think that this assessment of the situation is mistaken.

<sup>21</sup> Wright does not introduce this strategy in an effort to defend SE; in the context in which he proposes the strategy, he curiously just assumes the truth of SE and aims to use the strategy to disarm Cartesian-style modal arguments against the identity theory. But given that SE does not apply to a variety of philosophically important topics (e.g., logic, mathematics), as discussed in §6.2, the applicability of SE to any particular subject is not settled in advance. Rather, it must be settled on a case-by-case basis, for SE fails to apply to a great diversity of a posteriori correlations (see note 32 for examples). Therefore, Wright may not base his defense of the a posteriori identity theory on SE without eliciting intuitions showing that SE applies to mental properties. The same dialectical point applies mutatis mutandis to Levin and, evidently, Nagel.

<sup>22</sup> Cf. Bealer, "Mental Properties".

206 / **George Bealer**

To repeat, Kripke holds that there is a genuine conflict between his thesis that, say, it is necessary that Hesperus = Phosphorus and the ordinary assertion that it could have turned out that Hesperus was not Phosphorus.<sup>23</sup> Kripke takes there to be a conflict because he believes that ‘[I]t could have turned out that p entails that p could have been the case’.<sup>24</sup> And he believes that, if conflicts like this cannot be resolved, his argument for SE would be foiled. His resolution is to hold, that all, or most, of our intuitions are correct and that the apparent conflict among our intuitions is only an illusion resulting from the fact that the sort of prima facie anti-SE intuitions we have been discussing are *misreported*. The inaccurate statement ‘It could have turned out that p’ is accurately stated thus: ‘It is possible that a population of speakers in an epistemic situation qualitatively identical to ours would make a true statement by uttering “p” with normal literal intent’.

But this sort of metalinguistic rephrasal is untenable because of familiar problems concerning fine-grained intensional content. For example, it runs afoul of the Langford-Church translation test. Church describes this test thus: “[W]e may bring out more sharply the inadequacy of [an analysis] by translating into another language . . . and observing that the two translated statements would obviously convey different meanings to [a speaker of the other language] (whom we may suppose to have no knowledge of English).”<sup>25</sup>

Likewise, Kripke’s rephrasal runs afoul of analogues of the sorts of considerations raised by Tyler Burge and Stephen Schiffer against metalinguistic rephrasals of propositional-attitude reports.<sup>26</sup> For instance, Burge raises several objections (pp. 94–9) to metalinguistic rephrasals of the belief that his arthritis man forms when the doctor tells him, “You cannot have arthritis in the thigh.” Here are three examples. First, like Church, Burge objects that the metalinguistic reformulations prevent the relevant beliefs from being shared across language communities (p. 96). Second, upon hearing what the doctor told him, arthritis man forms a belief (naively, the belief that he cannot have arthritis in the thigh) that results in great relief—indeed, a dissipation of his fears (p. 95). But this relief is plainly not produced by a belief about the semantics of English (concerning the reference of the English word ‘arthritis’). Third,

<sup>23</sup> Kripke, *Naming and Necessity*, 103–5, 140–4.      <sup>24</sup> *Ibid.* 141–2.

<sup>25</sup> Alonzo Church, “On Carnap’s Analysis of Statements of Assertion and Belief”, *Analysis*, 10 (1950), 98.

<sup>26</sup> Burge, “Individualism & the Mental”. Stephen Schiffer, *Remnants of Meaning* (Cambridge, MA: MIT Press, 1987).

**Intuition and Modal Error / 207**

and relatedly, when *occurrent* mental events are at issue, arthritis man “may be brought up short by a metalinguistic formulation of his just-completed ruminations, and may insist that he was not interested in labels” (p. 97). Schiffer raises additional considerations. For example, he complains that the specialized semantical concepts required for satisfactory metalinguistic paraphrases are simply too sophisticated to enter into the contents of an ordinary person’s beliefs (p. 68). These objections clearly generalize to Kripke’s metalinguistic rephrasal strategy.

There is, however, an extremely simple alternative assessment of the situation, which now seems to me to be correct. Kripke held that there is no conflict in the intuitions at issue, but there is a conflict in the reports of those intuitions. The alternative response is simply to deny that there is conflict even in the reports. When we say (in the relevant situation) that it could have turned out that Hesperus was not Phosphorus, we are simply not contradicting the SE thesis that it is necessary that Hesperus = Phosphorus. Why? Because we are just employing an established epistemic use of ‘could’, namely, the ‘could’-of-qualitative-evidential-neutrality. As we saw in §2, this use of ‘could’ simply does not collide with the metaphysical use. End of story. Kripke took there to be a conflict in the *reports* of our prima facie pro- and anti-SE intuitions (insofar as he believed that “*it could have turned out that p* entails that *p* could have been the case”). True enough—there is a conflict when the uses of ‘could’ are the same. That is, ‘It could have turned out that Hesperus  $\neq$  Phosphorus’ and ‘It could not be the case that Hesperus  $\neq$  Phosphorus’ are outright contradictory when ‘could’ is used the same way in both sentences. But in the context of Kripke’s discussion his first and second uses of ‘could’ are simply not the same. So goes the alternative assessment. As soon as we see this, the appearance of conflict between the reports (as well as the intuitions expressed by them) vanishes.<sup>27</sup>

Remember that at this stage of the dialectic we are concerned with those who do in fact have the relevant pro-SE intuitions and who are concerned with what to do about the prima facie conflict between these intuitions and the apparently anti-SE intuitions. A philosopher in this position might hold that my alternative assessment of the apparent conflict is all well and good, but nevertheless wonder how we are to be sure whether we have

<sup>27</sup> Thus, just as in the case of Galileo’s Paradox, Kripke’s present puzzle arises from failing to notice a relevant distinction: as soon as we see it, we see that there was never a contradiction.

208 / **George Bealer**

an intuition of a metaphysical possibility or an epistemic possibility, say, qualitative epistemic neutrality. In “Modal Epistemology and the Rationalist Renaissance,” I proposed a test for exactly this purpose.

Suppose you are considering the modal status of one of the classic hypothetical-case propositions  $p$  (e.g., that Aristotle was not the teacher of Alexander, that Phosphorus is not visible in the morning, that there is water on twin earth). Typically, for the philosophical purpose at hand, you may bypass the question of whether  $p$  is metaphysically possible and consider instead whether  $p$  is a contingent proposition. Therefore, one may without loss recast our question in that idiom. And since a survey of cases shows that ‘contingent’—unlike ‘possible’, ‘could’, and so forth—does not have an epistemic reading, one need no longer worry about the possibility of equivocation. Consequently, concerning the twin-earth case, philosophers who share Kripke’s apparently pro-SE intuitions may shift their attention to the intuition that the twin-earth proposition is a contingent proposition—that is, the intuition that it is contingent whether there is a twin earth macroscopically, but not microscopically, like earth where there is a water-like stuff which is not genuine water.<sup>28</sup> But they may not do this for the epistemic possibility that water could have turned out not to be  $H_2O$ . The sort of philosopher that we are presently concerned with simply lacks the intuition that the proposition that water is not  $H_2O$  is a contingent proposition.

Thus we find that our first potential source of modal error—the alleged confusion of epistemic and metaphysical possibility—is relatively easy to untangle. In fact, in its most famous instances, it was not a source of error at all. The only error in those instances was in thinking that any of the relevant modal intuitions (either the apparently anti-SE or the pro-SE intuition) must have been erroneous in the first place.

## 6.2 Semantic Stability

For a certain important class of propositions, apparently anti-SE intuitions cannot be disarmed in the way just discussed. I have in mind propositions that are *semantically stable*.<sup>29</sup> Very roughly, a proposition  $p$  is semantically stable iff

<sup>28</sup> Or the intuition that it is contingent whether there is a twin earth macroscopically but not microscopically like earth where there is a water-like stuff and, *if* all actual samples of water are  $H_2O$ , *then* these twin-earth samples are not genuine water.

<sup>29</sup> I first introduced the notion of semantic stability, together with the following account, in “Mental Properties”.

**Intuition and Modal Error / 209**

no proposition besides  $p$  itself can play exactly the same cognitive role for our twin-earth counterparts as  $p$  plays for us. More precisely,

For thinkable  $p$ ,  $p$  is semantically stable iff, necessarily, if  $p$  plays some cognitive role in the mental life of a community  $c$ , then it is necessary that, for any other community  $c'$  in qualitatively the same epistemic situation as  $c$ , no proposition can play that role other than  $p$  itself.

Expressions have (or lack) a corresponding sort of epistemic invariance.<sup>30</sup> Natural kind terms are paradigmatic semantically unstable terms—‘water’, ‘gold’, ‘heat’, ‘beech’, ‘elm’, etc. By contrast, the core vocabulary of the a priori disciplines—logic, mathematics, philosophy—is semantically stable: ‘some’, ‘all’, ‘and’, ‘if’, ‘is identical to’, ‘is’, ‘necessarily’, ‘possibly’, ‘true’, ‘valid’, ‘0’, ‘1’, ‘+’, ‘÷’, ‘property’, ‘quality’, ‘quantity’, ‘relation’, ‘proposition’, ‘state of affairs’, ‘object’, ‘category’, ‘conscious’, ‘sensation’, ‘pleasure’, ‘pain’, ‘emotion’, ‘think’, ‘believe’, ‘desire’, ‘decide’, ‘know’, ‘reason’, ‘evidence’, ‘justify’, ‘understand’, ‘explain’, ‘purpose’, ‘good’, ‘fair’, ‘ought’.

A successful argument that SE holds in a particular case (e.g., necessarily,  $\text{water} = \text{H}_2\text{O}$ ) consists of two steps.<sup>31</sup> First, pro-SE intuitions are elicited supporting the thesis that the proposition at issue is an a posteriori necessity: in all known cases, these intuitions either are or can be reworked into twin-earth style intuitions. Second, one shows that prima facie anti-SE intuitions (e.g., the intuition that it could have turned out either way whether  $\text{water} = \text{H}_2\text{O}$ ) can be disarmed but that, when anti-scientific-essentialists attempt to do the same for prima facie pro-SE intuitions (i.e., the intuitions elicited in step one), they fail. Note, however, that, although this two-stage defense of SE succeeds for natural-kind identities like ‘ $\text{Water} = \text{H}_2\text{O}$ ’, both steps fail in the case of semantically stable terms.

Let  $t$  be a semantically stable term. In connection with the first step, consider the  $t$ -analogue of the twin-earth argument for ‘water’. We are to contemplate the possibility of another planet (or world) macroscopically like earth (the

<sup>30</sup> Although, for ease of expression, I will take the liberty to pass back and forth between the metalinguistic notion and the objectival notion, the objectival notion is basic and the metalinguistic notion is a derived notion defined in terms of it. Identifying the objectival notion as basic allows the stable/unstable distinction to be uniformly and economically extended to all categories of entity. In addition, it allows us to consider situations in which there are conscious beings but no natural languages. And it requires no caveat about ambiguity; after all propositions, concepts, properties, and so forth are not ambiguous.

<sup>31</sup> Here and elsewhere I will suppress existence conditions when they are not pertinent.

210 / **George Bealer**

actual world) but microscopically different. We are to consider the items *here* to which *t* applies, and we are then to ask whether, intuitively, *t* would fail to apply to the corresponding items *there* (which, by hypothesis, are microscopically different). The question is outlandish if *t* is a core term from logic, mathematics, or metaphysics (see the examples given above). Take, for instance, ‘property’. There are properties here; could there fail to be properties there?!

The same thing holds if *t* is a core term concept from epistemology, philosophy of mind, ethics, and so forth. Consider, for example, the concept of thinking. First, we are to suppose that on earth all and only thinking beings have a certain microstructure say, “T-fibers” (which are composed ultimately of hydrogen, oxygen, carbon, etc.). Consider a twin earth on which our Doppelgängers display “thinking”-behavior exactly like ours. It turns out, however, that, whereas our thinking—and our associated “thinking”-behavior—co-occurs with firing T-fibers, the “thinking”-behavior of our Doppelgängers co-occurs instead with firing  $T_{te}$ -fibers (composed ultimately of X, Y, Z, etc.). Would we say that these creatures are thinking? To be sure, we would not be *certain* that they are; macroscopic behavioral criteria never entail that a mental predicate applies. Nevertheless, *it would not be counterintuitive to say that they are thinking*. Note the contrast with water. It would be counterintuitive to say that samples of XYZ on twin earth are genuine samples of water. This intuition is the essential first step of the SE argument concerning ‘water’. The analogous intuition concerning ‘thinking’ is simply *missing*! Accordingly, the essential first step of the argument that SE applies to ‘thinking’ cannot even get off the ground. The general thesis is that this essential first step in the argument for SE fails for semantically stable terms.<sup>32</sup>

<sup>32</sup> Someone might think that even though this essential step in the argument for SE fails, considerations of simplicity create a presumption in favor of SE. But this is not so.. Even if in the actual world the atmosphere of every planet were chemically just like earth’s, we would not say that in a nonactual world a twin-earth would have no atmosphere if it were enveloped by *pure* oxygen. It would be preposterous to hold that simplicity creates a presumption in favor of restricting the term ‘atmosphere’ to the familiar mixture of gases on earth. Since there is no presumption in favor of SE for this and a large, remarkably diverse family of other, *prima facie* functional terms (e.g., soil, shelter, fuel, drink, and so forth), uniformity supports the conclusion that there is no presumption in favor of SE in the case of other *prima facie* functional terms like ‘thinking’—unless some reason independent of simplicity (e.g., genuine evidence or outright argument) can be given for taking ‘thinking’ to be different. For a detailed discussion (and additional examples), see Bealer “Philosophical Limits of Scientific Essentialism”.

**Intuition and Modal Error / 211**

Now for the second step in the SE argument. When we try to disarm the prima facie anti-SE intuitions concerning semantically unstable propositions (e.g., that it could have turned out that water  $\neq$  H<sub>2</sub>O) as we did in §6.1, we succeed. But in the case of analogous prima facie anti-SE intuitions concerning semantically stable propositions, this strategy always fails.

Though this may be seen on a case by case basis, it is also a consequence of the following general principle:

For semantically stable propositions  $p$ , if it could (epistemically) have turned out that  $p$ , it is metaphysically possible that  $p$ .

In slogan form, for semantically stable propositions  $p$ : epistemic possibility entails metaphysical possibility. (In symbols,  $\Diamond_{q-e-n}p \rightarrow \Diamond p$ .) This principle follows directly from (i) the definition of semantic stability and (ii) the truth conditions (given in §2.3; see also §§5–6.1) for the epistemic use of ‘it could have turned out’ needed for the defense of SE.<sup>33</sup>

Thus, when it comes to semantically stable propositions, our epistemic-possibility intuitions are a good guide to metaphysical possibility. So, if you are in doubt about the reliability of metaphysical-possibility intuitions but not the associated epistemic-possibility intuitions, the latter may without loss play the role formerly played by the former. The chance that we might be mistaking intuitions of epistemic possibility for intuitions of metaphysical possibility poses no threat: even if we are, no error would result since intuitions of the epistemic possibility entail the corresponding metaphysical possibility for semantically stable propositions.<sup>34</sup>

Clearly, then, the standard routine for disarming prima facie anti-SE intuitions fails in the case of our semantically stable prima facie anti-SE

<sup>33</sup> This principle is the basis for the argument against the identity theory I give in “Mental Properties”. Some people mistakenly believe that Kripke articulates this principle and indeed that his argument against the identity theory in *Naming and Necessity* is based on it; but examination of the text shows that this is plainly not so. See note 17 above.

<sup>34</sup> All the same points hold for “mixed” intuitions of the sort discussed above. All we need to do is replace the semantically unstable component with a necessary condition of this semantically unstable element that is semantically stable. As long as the substitute semantically stable concept is sufficiently specific, we will have the intuition that the new proposition is epistemically possible if we had the intuition that the original proposition was. And, again, the epistemic possibility would entail the metaphysical possibility.

212 / **George Bealer**

intuitions. The upshot is that neither stage of the argument for SE goes through for semantically stable concepts and propositions.<sup>35</sup>

This brings us back to the main thread of our discussion of modal error. At the close of §4 we isolated two sources of modal error with special bearing on SE: the first concerned the failure to distinguish between metaphysical and epistemic possibility; the second, the local misunderstanding of one's concepts. In §5 we began our investigation of alleged modal errors having to do with the first source. In §6.1 we completed this investigation, showing that the supposed instances of such modal errors identified by Kripke in his defense of SE simply were not genuine errors at all. In the course of §6.2 we considered a reconstruction of the two-step argument for SE as it applies to the standard class of semantically unstable propositions (that water = H<sub>2</sub>O; etc.). This reconstruction enables us to see more clearly where the resolution given in §6.1 fits into the larger argument for SE (namely, in step two).

Our discussion of the first source of modal error focused on people, like Kripke himself, who have pro-SE modal intuitions but who also report having anti-SE modal intuitions. By contrast, our discussion of the second source of modal error will focus on people who have anti-SE modal intuitions but simply lack pro-SE modal intuitions. We now turn to this discussion.

— ⊕ —

## 7. Modal Error and Local Misunderstanding

Stephen Yablo presents an account of modal error—specifically, how anti-SE modal intuitions can be in error.<sup>36</sup> Yablo is not concerned with errors resulting from conceptual illusions, limitations on intelligence, inattentiveness, and so

<sup>35</sup> Scientific essentialism therefore provides us with no reason to doubt the possibility of a priori knowledge of semantically stable modal truths. What could block this possibility? Evidently, nothing could unless there are necessary cognitive limitations, notably on intelligence. If correct, these considerations suggest the following, which I call the *Semantic Stability Principle*: For knowable semantically stable propositions *p*, if *p* is necessary, *p* is knowable a priori. Given that the central truths of the a priori disciplines are semantically stable necessities, the Semantic Stability Principle implies the more circumscribed principle that the knowable central truths of the a priori disciplines are knowable a priori. This is just the traditional rationalist thesis of the Autonomy of the A Priori Disciplines.

<sup>36</sup> Yablo, "Is Conceivability a Guide to Possibility?" Yablo's discussion is stated in the idiom of 'conceivability' and 'inconceivability'; I will be reformulating it in what follows in the idiom of possibility and impossibility intuitions, whose use we defended in §2.1.

**Intuition and Modal Error / 213**

forth. Nor is he concerned with the above problem (on which Kripke spent so much time)—namely, the problem of reconciling our metaphysical and epistemic intuitions, which (I have argued) were correctly reported in the first place and were never in conflict. Yablo’s underlying concern is rather with full-fledged errors in intuitions about metaphysical possibility.

Yablo holds that these errors have two potential sources, in each case mistaken *beliefs*:

- (a) mistaken a posteriori beliefs (e.g., someone who mistakenly believes that Hesperus  $\neq$  Phosphorus might have the intuition that Hesperus could outlast Phosphorus), or
- (b) mistaken beliefs regarding the relationship between such a posteriori beliefs and associated modal truths (someone might deny that, if Hesperus = Phosphorus, then necessarily Hesperus cannot outlast Phosphorus).

I am here less interested in class (a), for practiced dialecticians have the ability to proceed using exclusively “pure” a priori intuitions, namely, those that survive even under the hypothesis that such a posteriori beliefs (both pro- and con-) are unjustified or mistaken.<sup>37</sup>

How do people come to have erroneous modal intuitions belonging to the second class? Yablo’s answer is that they are somehow produced by underlying class (b) beliefs, which by hypothesis are false. But it is plausible that at least some people have class (b) beliefs that are based, ultimately, on relevant concrete-case intuitions. But since such a person’s class (b) beliefs are by hypothesis false, presumably a number of the intuitions upon which these false beliefs are based must themselves be false. What explains why *these* intuitions go wrong? If the explanation is that they too are produced by false class (b) beliefs, we go round in a circle. It seems, therefore, that we need something besides, or at least in addition to, Yablo’s belief-based explanation of class (b) intuition errors.

An example will help to bring out the same point, but in another way. Suppose two empirically well-informed, dialectically skilled philosophers have conflicting concrete-case SE intuitions. For example, suppose that, upon first

<sup>37</sup> In fact, by exercising this ability in the context of pure a priori philosophizing, one’s natural-kind intuitions will actually diminish in number by virtue of this ability—for instance, against the background of the hypothesis that all and only water on earth is composed of H<sub>2</sub>O, the intuition that necessarily water contains hydrogen; and against the hypothesis that water samples on earth have a highly disuniform composition, the intuition that water could have lacked hydrogen—thereby all but eliminating disagreements of the sort associated with class (a).

214 / **George Bealer**

considering the twin-earth example, Hilary Putnam had the intuition that the samples of XYZ would not be water whereas Rudolph Carnap, upon hearing the same example, had the contrary intuition.<sup>38</sup> How are we to explain this?

According to Yablo's explanation, Putnam's and Carnap's beliefs—specifically, their class (a) and class (b) beliefs—must at the time in question be relevantly different. But surely this need not be so. Surely we may suppose that at the time in question both men were empirically well-informed and that both believed the prevailing anti-essentialism of the day (in particular, both believed that water = H<sub>2</sub>O and that this fact is only contingent). In spite of this, they have opposing intuitions about the twin-earth case. As it stands, Yablo's account does not explain why this occurs.

We are thus in need of something besides (or in addition to) Yablo's belief-based explanation of class (b) intuition errors. Our analysis of what it is to understand a concept might provide the missing pieces. The simplest explanation in this vein would be that either Putnam or Carnap outright *misunderstands* the concept of being water. In some cases, this is no doubt the right explanation, but surely not in the case of Hilary Putnam or Rudolph Carnap: these eminent philosophers of science did not misunderstand the everyday concept of being water!

A more plausible explanation is that, in spite of having a full underlying mastery of the concept, Carnap *locally misunderstands* it. That is, he has a local (i.e., in principle temporary) disruption of his otherwise full understanding of the concept. This is analogous to the sense in which a patient would be given a clean bill of health at his annual check-up, despite his having a cold at the time of the check-up—the patient's health is locally (temporarily) disrupted, but his standing health is impeccable.

An example might help to illustrate this phenomenon of local misunderstanding.<sup>39</sup> A student, musing about prime numbers, reports having the intuition that  $-3$  is a prime number, and from this he concludes that in the definition of prime number the domain is not restricted to natural numbers but includes all integers, negative as well as positive. Fortunately, he also has a firm intuition that primes are divisible only by themselves and one, and he

<sup>38</sup> We may in addition suppose that Carnap has the further intuition that it is contingent whether water contains hydrogen. In view of our "contingency test" from §6.1, this would serve to show that his initial intuition concerns metaphysical possibility not some kind of epistemic possibility.

<sup>39</sup> This example is based on an exchange that occurred in one of my classes.

**Intuition and Modal Error / 215**

has the intuition that every negative integer,  $-n$ , is the product of itself and the number one and is also the product of  $n$  and  $-1$ . From this he rightly infers that negative integers cannot be prime. Then, he has the intuition that 3 is prime but that  $3 = 1 \times 3$  and  $3 = -1 \times -3$ . From this and the conclusion he just reached, he then infers that only natural numbers are permitted in the definition of prime. In view of this performance, the student plainly *understood* the concept of being prime all along. What went wrong early on was that he suffered a *local* lapse in his understanding of his concept of a prime number. Using only an a priori process, however, he was able to correct this lapse *on his own*, therein manifesting his underlying mastery of the concept.

Carnap is evidently in a somewhat similar situation. Not only does he have the intuition that, on twin earth, samples of XYZ would be water, he also would (if asked) have the corresponding mistaken categorial intuition that water is a macroscopic stuff (individuated by its macroscopic properties)—as opposed to a compositional stuff (individuated by its composition). But this categorial misunderstanding is (we may suppose) only *local*: it is correctable by Carnap *on his own* (without the aid of any auxiliary empirical information) using the a priori (dialectical) process, specifically, by careful examination of further cases, say, *other* sorts of twin-earth cases (e.g., the diamond/cubic zirconium twin-earth case<sup>40</sup>), and by systematization of the results. That is, left entirely to his own a priori devices, Carnap would in the fullness of time become a scientific essentialist. (So too, given the patient's standing good health, his infection will in the fullness of time be cured by his own bodily resources, without any external intervention.)

Having identified the phenomenon of local misunderstanding, we are now in a position to complete a previously missing step in our dialectic concerning SE. For, in addition to resolving (or dissolving) the apparent conflict among our own intuitions, a full defense of SE requires assuring ourselves that we are not subject to relevant local misunderstandings. The line of reasoning given above shows how we are to go about establishing that there is no such local misunderstanding underlying our pro-SE intuitions.

The general point is that, at least in a large family of cases, the quality of one's understanding of one's concepts holds the key, not only to the correctness

<sup>40</sup> The diamond-appearing samples on twin earth are samples of cubic zirconium (the comparatively cheap material from which fake diamonds are commonly made on earth). Would Carnap really have had the intuition that those samples are diamonds?! Certainly not.

216 / **George Bealer**

of one's intuitions, but also to their incorrectness; furthermore, whether or not that understanding has only lapsed locally is the key to whether or not it is correctable a priori. And this also explains a related, historically important phenomenon. For, presumably, if Carnap has the intuition that in the twin-earth example the samples of XYZ would be water, he would likewise have the intuition that the proposition that puddles of water contain hydrogen is a contingent proposition. In fact, just about everyone prior to the advent of SE had this intuition. The source of this remarkably widespread modal error is the very same local misunderstanding as in the Carnap example.

Local misunderstanding was of course a primary target of Socratic elenchus and is responsible for especially recalcitrant instances of Plato's problem of *doxa*. The phenomenon is ubiquitous in philosophy, and for that reason has significant implications for philosophical method. Amongst other things, before one is entitled to declare a philosophical conclusion final, one must always first assure oneself that one has not been the victim of local misunderstanding. And it should be borne in mind that local misunderstanding can be deeply hidden so that the dialectical recovery is long and hard. But ultimately it is a great boon to philosophy, for it allows us to tolerate colliding intuitions without having to abandon our classical method of intuition-driven philosophical investigation.

The more immediate moral of this discussion is thus that, besides Yablo's class (a) and class (b) belief-based intuitional errors (on the supposition that Yablo is right about these two sources of local error), there are two other classes:

- (c) those resulting from local misunderstanding, and
- (d) those resulting from out-and-out misunderstanding (for example, the sort of modal error Burge's arthritis man would be guilty of, namely, intuiting that it is possible to arthritis in the thigh).

Of course, analogous conclusions hold for the phenomenon of local lapses in the completeness (vs. correctness) of one's understanding of one's concepts.

## 8. Skeptical Accounts of Modal Error

We have thus far been considering a model of modal error according to which intuitions retain their evidential force, and according to which errors can

**Intuition and Modal Error / 217**

in principle be identified and eliminated by subjecting our intuitions to the sort of a priori dialectic sketched above. I mentioned at the start of §5 that certain proponents of SE endorse radical accounts of modal error according to which large families of our modal intuitions are systematically unreliable. In the present section, I will assess a typical, but admirably careful and well-articulated, example—namely, Christopher Hill’s account.<sup>41</sup> I will show that this account is ultimately inadequate: it attributes intuitions to a great many people that they simply do not have, and, moreover, it would undermine the evidential basis of SE, leaving proponents of SE in an epistemically self-defeating position. It is worth going into some detail on this matter since all radical accounts fall into similar self-defeat.<sup>42</sup>

Hill is interested in a family of intuitions of the form: it is possible for something to be A and not-B (where, for example, A is a common-sense kind and B is a theoretical kind). Such intuitions are generated as follows:

[T]here is a class H of psychological mechanisms whose members work as follows: where M is any member of H, M takes two concepts as inputs, and then, provided that it is possible to conjoin each of the concepts with the negation of the other without generating an inconsistency, and provided also that there is no available a posteriori reason to think that the two concepts are necessarily coextensive, M delivers an intuition of possibility as an output.<sup>43</sup>

This is supposed to be a perfectly general account of the mechanism responsible for generating possibility intuitions of the indicated form and, as such, is an adequate account only if it correctly predicts (among other things) which metaphysical possibility intuitions one will have (or lack) concerning commonsense kinds and associated microstructural properties upon consideration of the relevant cases.<sup>44</sup>

<sup>41</sup> Hill, “Imaginability, Conceivability, Possibility and the Mind-Body Problem”, 65–72.

<sup>42</sup> For example, Levin, “Tortuous Dualism” and Wright, “The Conceivability of Naturalism”. For more on the topic of epistemic self-defeat, see Bealer, “The Philosophical Limits of Scientific Essentialism” and “The Incoherence of Empiricism”.

<sup>43</sup> Hill, “Imaginability, Conceivability, Possibility and the Mind-Body Problem”, 76.

<sup>44</sup> There is an obvious problem with Hill’s account as stated that can be avoided by revising the account so that ‘the two concepts are necessarily coextensive’ is replaced with ‘necessarily, the first concept’s extension includes the second concept’s extension’. Let us suppose that at relevant points this repair is in place.

218 / **George Bealer**

Various proponents of SE reasonably believe that SE extends to certain commonsense kinds even though they have no specific knowledge of the relevant a posteriori facts (e.g., of their chemical composition). For example, suppose that Saul Kripke (or any other scientific essentialist, e.g., Hill himself) believes that SE extends to various kinds of gemstones—say, rubies—even though he has no specific knowledge of the chemical composition of rubies.<sup>45</sup> In this case, surely Kripke would not have the intuition that it is metaphysically possible that there be rubies containing no aluminum oxide (silicon dioxide, aluminum silicate, or some such). Likewise, he would not have the intuition that it is metaphysically possible that there be rubies that *are* composed of aluminum oxide (etc.). He would simply not have such intuitions one way or the other (see note 37 above). But Hill's account entails that, in the envisaged circumstance, Kripke *would* have the intuition that it is metaphysically possible that there be rubies that contain no aluminum oxide (etc.), for (i) there is no inconsistency (or analytic impossibility, as Hill requires in other passages) in the conjunction of the concept of being a ruby and the negation of the concept of containing aluminum oxide and (ii) given that (by hypothesis) Kripke lacks the relevant empirical knowledge concerning the chemical make-up of rubies, he lacks "a posteriori reason to think that the two concepts are necessarily coextensive." Thus, Hill's account predicts the wrong result. (Of course, Kripke would no doubt have intuitions that it is *epistemically* possible, in one or more of the senses catalogued above, that there be rubies containing no aluminum oxide; but according to Hill's account he should be having the corresponding *metaphysical*-possibility intuitions as well.)

Besides wrongly predicting in this way the modal intuitions people will have, Hill's account lands proponents of SE in an epistemically self-defeating position. Let me explain. After spelling out the mechanisms just discussed, Hill goes on to reject the outputs of a certain subcategory of these mechanisms as highly unreliable:

<sup>45</sup> LaPorte ("Chemical Kind Term Reference and the Discovery of Essence", *Noûs*, 30 (1996), 122–3) reports that mineralogists do not consider *all* minerals that share the chemical composition of rubies to be rubies. This, however, does not appear to collide with the SE claim that a *necessary* condition of a mineral's being a ruby is that the mineral have the chemical composition that rubies in fact have. If I am wrong about this, just choose another relevant a posteriori necessity.

**Intuition and Modal Error / 219**

[W]e are committed to holding that if M is a mechanism that produces conceivability-based intuitions to the effect that a *physical* commonsense kind is separable from its correlated theoretical kind, then that mechanism is highly unreliable.<sup>46</sup>

But how does Hill know that this mechanism is highly unreliable? Because its outputs conflict with SE. For example, such mechanisms would (for people who lack the a posteriori knowledge of the chemical composition of water samples on earth) take as inputs the concept of being composed of water and the concept of containing H<sub>2</sub>O and deliver as output the intuition that it is possible for there to be something composed of water but not containing H<sub>2</sub>O. But how does he (or we) know that SE is correct (assuming that the mechanism is unreliable)? Herein lies the self-defeat: for the justification of SE relies on intuitions which, according to this account, are of the sort generated by this highly unreliable mechanism. Thus, the account renders SE unjustifiable—and so too Hill’s thesis that the intuitions delivered by the mechanism are unreliable.

To see why, note that there are two ways one might come to know that it is metaphysically necessary that water contains H<sub>2</sub>O:

- (i) One might follow Kripke and Putnam’s method and try to establish it *indirectly* by means of concrete-case intuitions such as twin earth intuitions (cf. step one in §6.2).
- (ii) One might bypass such concrete-case intuitions and instead try to establish it *directly* by means of a (theoretical) intuition that it is impossible for there to be water that contains no H<sub>2</sub>O (or that it is metaphysically necessary that water contains H<sub>2</sub>O).

The dilemma for Hill is that his mechanism stands in the way of our coming to know these modal facts by either route.<sup>47</sup>

Consider the second route first. Of course, those who lack the empirical information that water and H<sub>2</sub>O are coextensive will lack a posteriori reason to think that the associated concepts are *necessarily* coextensive. So, according to Hill, the mechanism H will lead such people to have the intuition that possibly there is water that contains no H<sub>2</sub>O. Given Hill’s mechanism, to avoid having this intuition one must come to possess an a posteriori reason to think that the associated concepts are necessarily coextensive. As a first step, one must

<sup>46</sup> Hill, “Imaginability, Conceivability, Possibility and the Mind-Body Problem”, 78.

<sup>47</sup> Might there be a third route that relies solely on empirical evidence? No. As discussed below, empirical evidence alone cannot ever serve as evidence that something that is so is metaphysically necessary. See “The Philosophical Limits of Scientific Essentialism” for further discussion.

220 / George Bealer

acquire the empirical information that water and H<sub>2</sub>O are coextensive. But, on its own, this does not provide our person with a reason to think that *necessarily* water and H<sub>2</sub>O are coextensive. As Kant tells us, “Experience tells us, indeed, what is, but not that it must necessarily be so, and not otherwise.” No amount of empirical evidence (i.e., phenomenal experience and sense perception) can ever reveal whether something that is so is necessary; on the contrary, if we limit ourselves to empirical evidence, we are in the position of Quinean radical empiricists and, therefore, should conclude that there simply are no modal truths.<sup>48</sup> So, after learning that water and H<sub>2</sub>O are coextensive, how does one go on to know that they are necessarily coextensive?

Suppose that Hill answers that, upon learning that water and H<sub>2</sub>O are coextensive, one would straight off have the intuition that it is impossible that water and H<sub>2</sub>O not be coextensive and that this intuition would provide one with the requisite reason for thinking that they are necessarily coextensive. But this move is not available to Hill. To see why, recall that, according to Hill’s mechanism, after acquiring the empirical information but prior to having the intuition that it is impossible that water and H<sub>2</sub>O not be coextensive, one would (upon considering the question) have the intuition that it is possible that water and H<sub>2</sub>O are *not* coextensive. Given this, consider what would happen if one considered the question whether it is *possible* or, instead, *impossible* for water and H<sub>2</sub>O not to be coextensive. By the present supposition, one would have the intuition that this is *impossible*; and, at the same time, according to Hill’s mechanism, one would have the intuition that this is *possible*! That is, one would have simultaneous contradictory intuitions. But, in such an irrational state, one certainly would not have a reason to favor the impossibility intuition, and so one will not have acquired the missing reason for thinking that water and H<sub>2</sub>O are necessarily coextensive. If there is no alternate justificatory route to SE, skepticism would be inevitable.<sup>49</sup>

<sup>48</sup> Admitting testimony as evidence does not help the Quinean out of this problem, for how did our informants move from phenomenal experience and sense perception to knowledge of what is necessary? As Kripke says, “*Philosophical analysis* tells us that they [i.e., propositions of the type in question] cannot be contingently true, so any empirical knowledge of their truth is automatically empirical knowledge that they are necessary. This characterization applies, in particular, to the cases of identity statements and of essence.” (*Naming and Necessity*, 159, emphasis added.) And to know that identity statements and statements of essence cannot be contingently true, nonempirical resources are needed.

<sup>49</sup> Alternatively, suppose that Hill were to suggest that, once one has learned that water and H<sub>2</sub>O are coextensive, then (upon considering the question) one would straight off have the

**Intuition and Modal Error / 221**

Two final points about this horn of the dilemma. We have been focusing on the issue of commonsense kinds. But analogous considerations would hold if instead we were to focus on the question of whether the associated common nouns are rigid designators. Likewise, analogous considerations would hold if we focused on particulars (and their associated proper names), rather than kinds. There simply is no place Hill can get a foothold from which to justify SE.

Thus, the effort to justify SE via the *direct* route (i.e., bypassing concrete-case intuitions) leads to skepticism. (This is no surprise, for when it comes to theoretical issues already rife with controversy, we have no choice but to turn to our concrete-case intuitions.) This brings us to the other horn of the dilemma, on which one attempts to justify SE using concrete-case intuitions. Most of us came to conclude that it is necessary that water is coextensive with H<sub>2</sub>O by reflecting on concrete-case intuitions—for example, twin-earth intuitions. (This was Kripke’s route although Kripke himself does not frame his cases in terms of “twin earth” but rather in terms of possible worlds.<sup>50</sup>) In all known cases, the relevant concrete-case intuitions either are twin-earth intuitions or can be reworked into twin-earth style intuitions. But, once again, given Hill’s account of modal error, this justificatory route leads to failure. For, just as in the case of the direct route, Hill’s account implies that these intuitions are highly unreliable.

For instance, Putnam elicited the intuition that it is possible for the world to be macroscopically as it is in actuality even though the watery stuff is not

intuition that they are necessarily coextensive and that this intuition would provide one with the requisite reason for thinking that they are necessarily coextensive. In this case, the same conclusion would result. For, on the assumption that Hill’s mechanism M is operative, there would surely have to be a corresponding mechanism H’ that generates associated contingency intuitions (vs. possibility intuitions): specifically, whenever one considers the question whether this proposition is *necessary or contingent*, H’ would generate the intuition that the proposition that water and H<sub>2</sub>O are coextensive is contingent. But given our supposition on behalf of Hill, one would also have the intuition that this proposition is *necessary*. That is, one would have simultaneous contradictory intuitions. So, once again, the missing justification of SE is unavailable, and skepticism would result. The same conclusion holds if Hill were to focus on the conditional: if water and H<sub>2</sub>O are in fact coextensive, then they are necessarily coextensive. To see why, suppose someone were to intuit the truth of this conditional. Suppose further that the person knows that water and H<sub>2</sub>O are in fact coextensive and that the person considers the question whether this proposition (that water and H<sub>2</sub>O are coextensive) is necessary or contingent. Then, the person surely would have the intuition that the proposition that they are coextensive is necessary. Given this, the remainder of the argument in the text goes through just as before.

<sup>50</sup> Kripke, *Naming and Necessity*, 131–3.

222 / **George Bealer**

H<sub>2</sub>O (but rather XYZ). Putnam then elicited the intuition that samples of this watery stuff would not be water. Let us focus on the first of these two intuitions, which is a possibility intuition conjunctive in form. Consider the second conjunct, namely, that the watery stuff is not H<sub>2</sub>O. According to Hill, the intuition that it is possible that the watery stuff is not H<sub>2</sub>O is generated by a “highly unreliable” mechanism. (After all, watery stuff is a commonsense kind; H<sub>2</sub>O is a theoretical kind; there is no relevant analytic connection between watery stuff and H<sub>2</sub>O; and we are capable of lacking a posteriori evidence regarding the fact that the actual extension of the commonsense kind watery stuff coincides with the actual extension of the theoretical kind H<sub>2</sub>O.) Since this intuition is generated by a highly unreliable mechanism, it is not to be trusted. Now, if someone lacking the relevant a posteriori information has the intuition that it is possible that watery stuff not be H<sub>2</sub>O, in almost all instances that person would also have the twin-earth intuition (that it is possible for the world to be macroscopically as it is in actuality even though the watery stuff is not H<sub>2</sub>O). But if the former intuition is not to be trusted, neither is the twin-earth intuition. Since this evidently generalizes to pretty much all concrete-case intuitions used to justify SE, SE would lack the missing justification, if Hill’s account were correct.

Hill might respond that the watery stuff is not the sort of commonsense kind on which his unreliable mechanism operates; on the contrary, it operates only on genuine “natural kinds” (in some preferred sense). But this only takes us in a circle, for how are we to distinguish natural kinds (in this preferred sense) from other kinds? Presumably, the answer resides in the fact that natural kinds, unlike the other commonsense kinds, are subject to SE.<sup>51</sup> Again, however, the very intuitions required for distinguishing those kinds that are subject to SE from those that are not are generated by a “highly unreliable” mechanism on Hill’s account.

Hence, Hill is in a dialectically self-defeating position. By condemning as unreliable such a wide range of our modal intuitions, he has left himself with no evidential basis for the SE claims that he employs in his argument. This result generalizes to other skeptical accounts of modal error. (These points hold in spades for philosophers who would reject intuitions altogether, for

<sup>51</sup> In this connection, it is important to remember that there is a very wide range of commonsense kinds to which SE does not apply: for example, atmosphere, dirt, fuel, clothing, and so forth.

**Intuition and Modal Error / 223**

how do they propose even to refute logical behaviorism, instrumentalism, phenomenism, inductivism, and all the other philosophical dinosaurs that they no longer believe in?)

Clearly what one needs in order to justify SE is a more forgiving account of modal error according to which the sort of intuitions needed to establish SE are not condemned as erroneous. This is what I have tried to provide. I have catalogued various standard sources of intuitional error—cognitive deficiency, conflation of one's concepts, out-and-out misunderstanding of one's concepts, underdescription and context-sensitivity, and so forth. I then went on to isolate a largely overlooked source of modal error, which skeptical accounts do not accommodate, namely, the local misunderstanding of one's concepts. Fortunately, even when one's understanding of certain pivotal concepts has lapsed locally, one's larger body of intuitions (given suitably good cognitive conditions) would be sufficiently reliable to allow one to ferret out the modal errors resulting from this lapse in understanding by means of dialectic and/or a process of a priori reflection. As we have seen, this source of modal error, and the capacity to overcome it, has wide-ranging implications for philosophical method—including, in particular, its promise for disarming skepticism about the classical method of intuition-driven philosophical investigation itself.

