

Is Conceivability a Guide to Possibility?*

STEPHEN YABLO

University of Michigan, Ann Arbor
University of Toronto

...because I find absence of incompatibility, because, that is, I am without a certain perception, I am to call my idea compatible. On the ground of my sheer ignorance, in other words, I am to know that my idea is assimilated, and that, to a greater or lesser extent, it will survive in Reality.

F. H. Bradley, *Appearance and Reality*

I. Introduction

Some propositions are "possible": the way they represent things as being is a way things metaphysically *could* have been. Other propositions are not in this sense possible. How do we tell the difference? Or more particularly, of the possible propositions, how do we tell *that* they are possible?¹ Hume's famous answer is that it is

an establish'd maxim in metaphysics, *That whatever the mind clearly conceives, includes the idea of possible existence*, or in other words, *that nothing we imagine is absolutely impossible*.²

* Research for this paper was supported by the National Endowment for the Humanities and the Social Sciences and Humanities Research Council of Canada. For discussion and comments I am grateful to Paul Boghossian, Jim Brown, Jim Conant, John Devlin, Graeme Forbes, Hannah Ginsborg, Danny Goldstick, Sally Haslanger, David Hills, Eileen John, Gideon Rosen, Larry Sklar, William Taschek, David Velleman, Ken Walton, Nick White, Crispin Wright, Catherine Wright, and audiences at Davidson College, Queens University, Wayne State University, and Ohio State University.

¹ Sometimes, of course, this is easy. If a proposition *p* is true, and known to be, then its possibility can be inferred from *p* itself. The problem is to find grounds for thinking a proposition possible which is *not* known to be true, most obviously because it is false.

² Hume 1968, p. 32. The maxim seems to say that conceivability *suffices* for possibility. This is implausibly strong, so I propose to (mis)interpret Hume as claiming only

And if there is a seriously alternative basis for possibility theses, philosophers have not discovered it. So it is disappointing to realize that Hume puns on "establish'd." What the maxim *is*, is entrenched, perhaps even indispensable. But our *entitlement* to it has often been questioned.³

Doubts about Hume's maxim have a variety of historical sources. Some date back as far as Descartes's claim that, since he can conceive himself in a purely mental condition, his essence is only to think. "How does it follow," Arnauld asks, "from the fact that he is aware of nothing else belonging to his essence, that nothing else does in fact belong to it?"⁴ Others are as recent as the discovery by Kripke and Putnam of necessary truths knowable only *a posteriori*:

we can perfectly well imagine having experiences that would convince us...that water *isn't* H₂O. In that sense, it is conceivable that water isn't H₂O. It is conceivable but it isn't logically possible! Conceivability is no proof of logical possibility.⁵

Between times we find Reid and Kneale warning that if a proposition is true "for all you know," then you will find it conceivable whether it is possible or not. More than can be appreciated from a few examples, though, pessimism about conceivability methods has been a consistent theme in philosophy. When Mill says that

our capacity or incapacity of conceiving a thing has very little to do with the possibility of the thing in itself; but is in truth very much an affair of accident, and depends on the past history and habits of our own minds,⁶

he sums up the position of many authors and the instinctive assumption of many more.

Yet throughout this complicated history runs a certain schizophrenia in which, the theoretical worries forgotten, conceivability evidence is accepted without qualm or question. Hume's own famous applications of his maxim are a case in point. There is nothing necessary about the uniformity of nature, he says, for

We can at least conceive a change in the course of nature; which sufficiently proves, that such a change is not absolutely impossible.⁷

that the conceivable is *ordinarily* possible and that conceivability is *evidence* of possibility.

³ Arthur Pap writes that "there is no objection to the imaginability criterion simply because there is no alternative to it" (1958, p. 218). As the advice not to abandon a leaky lifeboat, this has its points. As factual observation, though—well, such objections are extremely common.

⁴ CSM II, p. 140.

⁵ Putnam 1975, p. 233. See Putnam 1990, pp. 55–57, for second thoughts.

⁶ Mill 1874, book II, chapter V, section 6.

⁷ Hume 1968, p. 89.

Causes are not strictly necessary for their effects, because the latter are conceivable as uncaused; nor are they sufficient since it is always conceivable that the effect should not ensue. Whatever our other differences with Hume, these arguments are normally credited with a good deal of persuasive force. Or consider a case from the philosophy of language. As everyone knows, 'Alexander's teacher' is not a rigid designator. How though does everyone know this? Well, we imagine a counterfactual situation in which Aristotle refuses Phillip's call, or dies of dysentery on the way to Macedonia. Such imaginings would be irrelevant to the rigidity of 'Alexander's teacher' if they gave no evidence of possibility.

In the actual conduct of modal inquiry, our theoretical scruples about conceivability evidence are routinely ignored. Double-think, though, is not the method of true philosophy. Those of us willing to be persuaded of *p*'s possibility by our ability to conceive it (and that is most of us, most of the time) should face the issue squarely: is this procedure ill-advised? There will be just one constraint on the discussion. Because the topic is not knowledge in general but knowledge of possibility, we will confine ourselves to problems or supposed problems *peculiar* to conceivability arguments. Such arguments have been charged, for instance, with trading on a confusion between two senses of 'could'; with implicit circularity; and with misclassifying most or all *a posteriori* impossibilities as possible.

Other, more sweeping, objections have also been raised. Two in particular deserve mention now, if only to put them aside for purposes of this paper. First is the traditional skeptical lament that

No independent evidence exists that conceivability is a guide to possibility—no evidence obtainable without reliance on the faculty under review.

True enough. But there is no independent evidence either that perception is reliable about actuality; and if the worst that can be said about conceivability evidence is that it is as bad as perceptual evidence, that may be taken as grounds for relief rather than alarm. Now though comes the objection from naturalism:

Granted the unavailability of any philosophically satisfying *reason* to think that perception is adequate to its task, we see at least how it *could* be. In fact perception itself brings word of sensory mechanisms seemingly hard at work monitoring external conditions. By contrast "we do not understand our own must-detecting faculty."⁸ Not only are we *aware* of no

⁸ Blackburn 1986, p. 119.

bodily mechanism attuned to reality's modal aspects, it is unclear how such a mechanism could work even in principle.⁹

Taken in a suitably flat-footed way, these claims are again true enough. But the same could be said about various other faculties, notably logical and mathematical intuition; and to judge by our reaction there, they constitute a reason less for mistrusting the faculty than for reconsidering either the nature of the target facts or the nature of our access to them.¹⁰

So much for the grand-scale objections. Ultimately they are going to require answers, but answers of a kind that the experience of philosophy has accustomed us to doing without. At any rate they are not the objections that concern me, or, I think, Arnauld, Reid, Kneale, etc. Two differences seem important. First, these philosophers seem prepared to bracket worries that arise also with other accredited ways of knowing, the better to focus in on what might be *especially* problematic about conceivability. Second, rather than simply deploring the *absence* of reason to think that conceivability *is* a guide to possibility, Arnauld and company offer *positive* evidence that it is *not* a guide. If the problem with conceivability methods was only that we could not prove, or explain, their reliability, then maybe we could live with that. But the problem is supposed to be that they are demonstrably *unreliable*.

II. Conceivability and the Modal-Appearance Test

What conceivability is is a question I hope to put off as long as possible. For now we can get by on a single assumption, one perhaps implicit in Hume's remark quoted above:

whatever the mind clearly conceives, includes the idea of possible existence, or in other words,...nothing we imagine is absolutely impossible.

As often when Hume takes himself to be saying the same thing twice, he seems here to be saying two quite *different* things:¹¹

- (a) what we imagine or conceive is *presented* as possible;
- (b) what we imagine or conceive *is* possible.

Where (b) claims for conceivability a certain *external* relation with possibility, (a) looks more like a partial *analysis* of conceivability, namely, that to

⁹ Cf. Wright 1986, pp. 206–7.

¹⁰ For a sense of the possibilities, see Coppock 1984, Forbes 1985, chapter 9, Bealer 1987, Sidelle 1989, and Yablo forthcoming.

¹¹ The classic example: “we may define a cause to be *an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second*. Or in other words *where, if the first object had not been, the second never had existed*” (Hume 1963, section VII, part II).

conceive or imagine that *p* is *ipso facto* to have it seem or appear to you that possibly, *p*. Without suggesting that Hume would go quite so far, I take the idea to be that conceiving is in a certain way analogous to *perceiving*. Just as someone who perceives that *p* enjoys the appearance that *p* is *true*, whoever finds *p* conceivable enjoys something worth describing as the appearance that it is *possible*.¹² In slogan form: *conceiving involves the appearance of possibility*.

Before trying to make the slogan clearer, let me stress that I am *not* tout-ing the “appearance of possibility” as all there is to conceiving,¹³ or the only thing conceiving can ever be. Far from trying to give the notion’s one true meaning, my aim right now is only to distinguish conceiving in the sense that matters from various *other* cognitive operations doing business under the same name. For as I will be interpreting it, the question whether conceivability is a guide to possibility concerns the kind of conceivability that *advertises* itself as such a guide. This means that if there are kinds of conceivability that do *not* portray *p* as possible—and there are—then for my purposes it *will not matter* if their modal guidance should prove unreliable.

Following in the tradition of Brentano, Husserl, and most recently Searle,¹⁴ suppose we take seriously the idea that many intentional states and acts—beliefs, desires, and perceptual experiences, for instance—have *satisfaction* conditions. And let us agree that these satisfaction conditions are at least in some cases the conditions under which the state in question is *true* or *veridical*. So, your belief that DeGaulle liked cheese is true just in case he did, and my perceptual impression that rain is falling is true just in case rain is falling.

From examples like these, one obvious conjecture would be that the *truth conditions* of an intentional state (assuming it has some) are a function of its *content*.¹⁵ But consider someone who, rather than *believing* that DeGaulle liked cheese, inwardly *denies* that he did. This person’s state has the same content as the believer’s, yet unlike the believer’s state it is correct just in

¹² Two notes about terminology. First, here and below I use ‘conceive that *p*’ and ‘find *p* conceivable’ essentially interchangeably. (But see note 59.) Second, ‘conceive’ has a *factive* sense—in which I don’t find *p* conceivable unless it is possible—and ‘perceive’ is *normally* factive—I don’t perceive that *p* unless *p*. In *this* paper, both terms are to be understood *nonfactively*. Thus ‘I perceived that *p* but it wasn’t true’ and ‘although I found *p* conceivable, it turned out to be impossible’ are perfectly in order. Out of order, though, will be the following: ‘I *veridically* perceived that *p*, but *p* wasn’t true’ and ‘although I *veridically* conceived that *p*, it turned out to be impossible.’

¹³ Later I’ll suggest that the conceiver enjoys this appearance *in a certain way*—by imagining a more or less determinate situation of which *p* is held to be a correct account.

¹⁴ See Dreyfus 1982, “Introduction” and *passim*; and Searle 1983.

¹⁵ Thus Searle: “To know the [representative content of an intentional state] is already to know [its satisfaction conditions], since the representative content gives us the conditions of satisfaction, under certain aspects, namely those under which they are represented” (Dreyfus 1982, p. 266).

case DeGaulle did *not* like cheese. So, *the truth conditions of an intentional state cannot be read off its content alone*; as the examples of denial, expectation and memory show, the state's psychological mode or manner is also relevant. This is crucial because *one* thing I will be taking "conceivability involves the appearance of possibility" to mean is that the truth conditions of an act of conceiving that *p* include, not the condition that *p*, as in perception, but the condition that *possibly p*. From now on I will express this by saying that *p*'s possibility *representatively appears* to the conceiver.

Maybe the analogy with perception can be carried a little further. Perceiving that *p* has in general the effect of *prima facie* justifying, to the subject, the belief that *p*, and thereby *prima facie* motivating that belief. Here the parenthetical "to the subject" is to indicate that the perceiver need only *feel* himself to be *prima facie* justified, that is, to cancel any suggestion that he is *prima facie* justified in fact. Thus someone convinced that he can judge sexual orientation at a glance might feel justified, on the basis of casual inspection, in believing a neighbor to be heterosexual, yet without possessing the slightest real evidence that this is so. That his neighbor is heterosexual epistemically appears to this person, even though his feeling of justification is quite misplaced. To have a word for this, let's say that *p* *epistemically* appears to me when some representative appearance I enjoy *prima facie* motivates me to believe that *p*, by making that belief seem to me *prima facie* justified.¹⁶

That our two readings of "appears" are compatible should be clear; the state that moves me to *believe* that rain is falling can surely be one with the *truth conditions* that rain is falling. Perhaps it could even be argued that the representative reading *entails* the epistemic one, for instance, that a visual experience with the truth conditions that *p* cannot help but move the experimenter to believe that *p*.¹⁷ However that may be, the readings are distinct, for the converse entailment fails: for me and I assume for others, it is only *epistemically* that the bull looks as though it is about to charge, or the car sounds like it's not going to make it through the winter.¹⁸ (Suppose that

¹⁶ For brevity, I'll speak simply of being *moved* to believe that *p*. (Why not define epistemic appearance in purely motivational terms? Because I do not want to say that *p* epistemically appears in cases where my motive for believing it is nonepistemic. Suppose I enjoy a representative appearance of someone offering to settle my debts if I will agree that *p*; this might tempt me to believe that *p*, but *p* does not epistemically appear to me.)

¹⁷ *Objection:* Someone confronted with the Muller-Lyer diagram enjoys the representative appearance that the lines are unequal, but unless the diagram is completely new to her, she does not *believe* that they are unequal. *Reply:* What epistemically appears to a subject turns not on her beliefs but on what she is *moved* to believe. And why speak of a Muller-Lyer *illusion* if typical observers aren't *moved* to believe the lines unequal?

¹⁸ Admittedly it is hard to draw a definite line between representative and (merely) epistemic appearances. Experts (matadors and mechanics) can enjoy representative appearances which to most of us are available only epistemically. But expertise is acquired gradually, and on the road to it there will be appearances not happily classified either

your car *does* make it through the winter. Then your experience has tempted you into a false belief, but it's not as though you were the victim of a sensory illusion!)

Back to the slogan "conceivability involves the appearance of possibility," should "appearance" here be taken in the representative sense or the epistemic one? *Both* senses are intended. Just as to perceive that *p* is to be in a state that (i) is veridical only if *p*, and that (ii) moves you to believe that *p*, to find *p* conceivable is to be in a state which (i) is veridical only if possibly *p*, and (ii) moves you to believe that *p* is possible.

With this background I can state my position. When we look at the standard objections to Hume's maxim, we find that they presuppose conceivability-notions that are neither mandatory nor particularly natural relative to the purposes at hand. Not natural, because none of them involves the *appearance* of possibility. Not mandatory, because there is an alternative notion, *philosophical* conceivability, that *does* involve this appearance and that sustains Hume's maxim against the objections. So the story has a negative part (sections III–IX) and also a positive one (sections X–XIV). At the end (section XV) I draw some tentative morals for the issue of realism vs. antirealism about modality.

III. The Confusion Objection

Strangely prevalent in philosophy is the idea that to find a proposition conceivable is to find that it *is true for all you know*. Since Reid explained conceiving *p* as "giving some degree of assent to it, however small,"¹⁹ the idea has been repeated by many authors; to choose a source almost at random, William Kneale says or implies that to find *p* conceivable is to "have in mind no information which formally excludes" that *p* is true.²⁰ Ignoring minor differences of formulation, suppose we let the proposal be that *p* is conceivable iff it is *not unbelievable*, or for short *believable*.²¹ (Remember that this is not to say that we see *p* as particularly *likely*, but just that we feel unable to rule it out.)

From an ordinary language perspective, the proposal is hard to argue with. Writing in the spring of 1990, Elizabeth Drew observed that German re-

way. For our purposes the indeterminacy doesn't matter; what will matter is the contrast between cases where *p* appears in both senses and those where it appears in neither.

¹⁹ Reid 1969, essay IV, chapter III. This isn't Reid's preferred account. Usually he says that to "conceive a proposition...is no more than to *understand distinctly its meaning*" (*loc. cit.*). Since one can distinctly understand the meaning of a contradiction, this is an obvious nonstarter as an analysis of the kind of conceivability which purports to discover possibilities. (For early discussion of the "some degree of assent" theory, see Mill 1874, book II, chapter V, section 6, and Mill 1868, vol. I, chapter VI.)

²⁰ Kneale 1949, p. 213.

²¹ Cf. Pap 1958, pp. 37–38, and van Cleve 1983, p. 37.

unification had “become conceivable only in the last few months.”²² Anyone reading this would take it to mean, not that our powers of imagination had suddenly improved, but that reunification could no longer be regarded as out of the question. Likewise if I call it inconceivable that there is a largest prime number, but conceivable that there is a largest *twin* prime, I am saying that although it is certain that the primes are infinite in number, with the twin primes, things are not so clear.

Suppose I find p conceivable in the sense of believable. Does this give me reason to think that p is metaphysically possible? In other words, do I acquire evidence in *favor* of a proposition’s possibility, by finding myself *without* evidence against its truth? That would be very strange, to say the least. Among other things it would have the result that there is a necessary limit on how bad my epistemological position can get: the *poorer* my evidence for p ’s truth, the *better* my evidence for its possibility.²³ (In the limit of perfect ignorance about p ’s truth, its possibility would be absolutely assured!) Yet the fact is that I can be *completely* in the dark about truth and possibility simultaneously, as for example with the twin prime conjecture.

Apart though from the sheer oddity of arguing from ignorance to substantive modal conclusions, how reliable are such arguments? Already in Reid we find the only plausible answer:

will it be said, that every proposition to which I can give any degree of assent is possible? This contradicts experience, and therefore [Hume’s] maxim cannot be true in this sense.²⁴

Reid doesn’t say what sort of “experience” he has in mind, but perhaps he was thinking of something he mentions later:

Mathematics afford[s] many instances of impossibilities in the nature of things, which no man would have believed if they had not been strictly demonstrated [that is, their impossibility would not have been believed if it had not been proved].²⁵

So propositions to which people *once* gave “some degree of assent,” say, the axioms of naive set theory, have often turned out *later* to be impossible. As an example of Kneale’s shows, it is not always necessary to wait. Speaking of Goldbach’s conjecture that every even number is obtainable as the sum of two primes, Kneale says that although it “looks like a theorem,...it may conceivably be false.”²⁶ Likewise it may conceivably be true. But if true, it is

²² *New Yorker*, March 19, 1990, p. 104. (At the time of writing reunification was far from a sure thing; to everyone’s surprise it occurred just a few months later.)

²³ Compare Bradley’s sarcastic remark that “merely because I do *not* find any relation between my idea and the Reality, I am to assert, upon this, that my idea is compatible.” The epigraph is in a similar vein: “On the ground of my sheer ignorance...” (Bradley 1969, pp. 345–46).

²⁴ Reid 1855, essay IV, chapter III.

²⁵ *Ibid.*

²⁶ Kneale 1949, p. 80.

necessarily true, and if false, necessarily false. Thus either the conjecture or its denial is a conceivable, that is to say a believable, impossibility. And the gimmick generalizes: we get a present-tense counterexample to the possibility of the believable *whenever* a proposition's truth-value is necessary but still unknown.

As a guide to possibility, then, conceivability *qua* believability is unreliable in the extreme. The fact that *p* might, for all I know, be true in the actual world, is just irrelevant to the issue whether it is true in some possible world or other. This leaves a puzzle, however: if the argument is as bad as that, why does there so much as *seem* to be an evidential connection? The answer is supposed to be that terms like "could" and "might" are *ambiguous*, which leads us into a certain confusion. Neglecting the distinction between what could be so in the sense that one is in no position to rule it out, and what could be so in the sense that it is metaphysically possible, we jump straight from the one to the other. According to the confusion objection, once this equivocation is noticed the appearance evaporates that conceivability argues for possibility.

IV. Believability

Without a doubt, sliding from epistemic to metaphysical "could" is something we sometimes do, though we really should not. But, could a mix-up this elementary²⁷ really be *all there is* to the conceivability maxim?

Probably the *locus classicus* of the supposed confusion is Descartes's argument in the *Meditations* for the possibility of disembodied existence. Finding in the "First Meditation" that there might, for all he knows, be no material things, he suggests in the "Second" that he can exist without them. Isn't Descartes reasoning here that since he "could" in the believability sense exist without benefit of matter, he "could" do it in the metaphysical sense as well?

Part of the problem with such an interpretation is just that the attributed argument is so awful. But never mind that: if Descartes is attracted to this sort of argument, why does he not use it more often? At this point in the *Meditations*, remember, Descartes finds virtually *everything* believable, including for instance that he is essentially a body, and that God does not exist. Shouldn't he then conclude that these other things are possible as well? To answer that he doesn't conclude that they are possible, because he doesn't *believe* that they are possible, treats Descartes as rather more arbitrary than his position requires. Surely it would be better if we could make him out to mean something *other* than "believable" by "conceivable," such that he does

²⁷ Among the many who have noticed it are Moore 1966, pp. 228ff; Sellars 1963, pp. 76ff; and Kripke 1980, p. 141.

not find it conceivable, in the sense *he* means, that he is essentially a body, or that God does not exist.²⁸

Or take the example of our finding it conceivable, in the sense of believable, both that Goldbach's conjecture holds and also that it fails. If the inference from epistemic "could" to metaphysical "could" were so inviting, then it ought to seem strange that not a single author has concluded that although in some possible worlds, every even number is the sum of two primes, in others one or more of them *stops* being the sum of two primes.²⁹ Was it just that they knew that in this case, such a conclusion would be counterintuitive? Again, a more sympathetic interpretation would be that conceivability, in the sense relevant to possibility, is a different thing from believability, and that neither Goldbach's conjecture nor its negation is conceivable in the relevant sense.

Earlier I agreed that "conceivable," as it occurs in daily conversation, usually *does* mean "believable." In fact more is true. As G. E. Moore noticed in an early paper,³⁰ not only "conceivable" but even "*possible*" normally indicates believability. Suppose, for example, that I tell you "it is possible that I was born on the moon." Assuming that I metaphysically *could* have been born on the moon, why does my statement sound so incredible? The reason is that "it is possible that *p*," where the embedded sentence is in the *indicative* mood, expresses uncertainty that *p* is false.³¹ Thus "it is possible that I was born on the moon" says, not that this *could* have happened although it didn't,

²⁸ Consider in this connection Michael Hooker's challenge to Descartes's argument: Existence in the absence of bodies is no more conceivable than existence in the absence of persons not identical to bodies. On his own principles, then, Descartes *could* have been identical to a body. But whatever is possibly a body is a body essentially; so, although Descartes's actual position is that he can exist without bodies, he could equally have concluded that he is essentially a body (my précis of Hooker 1978, section II).—But why think that Descartes finds it conceivable that he should have been identical to a body? The only evidence Hooker offers is that "he does not know at this point in his inquiry that there are any disembodied minds," and that if "reflective consideration...leads one to doubt that *p*, then the truth of not-*p* is at least conceivable" (p.181). However this is just to say that (reflective) believability suffices for cartesian conceivability, which is exactly what I deny. Hooker might counter that it is still mysterious why existing *as* a body should be any less conceivable than existing *without* bodies. Here is a suggestion: If all possible bodies are essentially bodies, and Descartes knows this, then to conceive himself identical to a body will be to imagine a world relative to which he is a body in *every* world. But how is Descartes to tell whether he can imagine a world like that without first attempting to imagine worlds in which he is *not* a body? Finding that he *can* imagine such worlds, Descartes is unable to conceive himself identical to a body. (Analogy: asked to think of a number such that all numbers are prime, you first consider whether you know of any *non*prime numbers. Realizing that you do, you find numbers of the first type unthinkable.)

²⁹ Compare Reid: "I have never found that any mathematician has attempted to prove a thing to be possible because it can be conceived..." (Reid 1969, essay IV, chapter III).

³⁰ "Certainty" (Moore 1966).

³¹ To a first approximation, anyway. See DeRose 1991 for a more sophisticated treatment.

but that I am not entirely convinced I was born on Earth. (To assert genuine possibility, I must use the subjunctive mood: "it is possible that I should have been born on the moon.")

None of this is really very interesting except as a reminder that philosophers sometimes use words differently from other people. In metaphysics, for example, "possible" is often used for something *other* than believability, and this whether the subjunctive mood is used or not. Mightn't something similar be true of "conceivable"? The view I called strangely prevalent above is not that "conceivable" *ever* means believable, but that this is what it *always* means, including in conceivability arguments. For the truth is that in conceivability arguments, or at least competent ones, "conceivable" rarely if ever means believable.

There are two directions to this: conceivable propositions need not be believable, and believable propositions need not be conceivable. The easy direction is the first. An old Jewish saying runs: "Life is so full of misery and woe; how much better it would have been never to have existed at all; yet how many of us are that lucky?" Thinking about this, I find it conceivable that I should never have existed. Never for a moment, though, do I find it believable that I *have* never existed. So here is an example of a conceivable proposition that isn't believable.³² Notice the underlying point: if conceivability entailed believability, then whenever one was *certain* that something was not the case, one would be unable to conceive it even as a possibility! This being absurd the entailment does not go through.

Of believable propositions that aren't conceivable, it is difficult to give a pure example, if this means a believable proposition which is positively *inconceivable*.³³ After all, if *p* is believable, then the *actual* world might for all I know be a *p*-world. So I am unlikely to have it appear to me that *p* cannot be true in *any* possible world.

Perhaps there can be an impure example though. Sometimes when we find ourselves unable to conceive a proposition, we don't find it inconceivable either; its modal status is *undecidable* on the available evidence.³⁴ Despite what you often hear, this is how it is with Goldbach's conjecture. No thought experiment that I, at any rate, can perform gives me the representational appearance of the conjecture as possible *or* as impossible, or the slightest temptation to believe *anything* about its modal character. So this is already an ex-

³² This gives, incidentally, another reason not to interpret Descartes as meaning "believable" by "conceivable." Probably there is nothing that Descartes finds more unbelievable than that he does not exist; yet for every created thing, Descartes finds it conceivable that it should not have existed. (Thanks to John Devlin for the next two sentences and the next note.)

³³ Although consider Tertullian: "Credo quia absurdum est."

³⁴ van Cleve 1983 distinguishes in a similar vein between *strong* and *weak* conceivability—"seeing" that *p* is possible vs. not "seeing" that it is impossible—and he describes Goldbach's conjecture as only weakly conceivable.

ample of a believable proposition that is not conceivable. But let me suggest some more interesting cases.

According to legend, the queen of Sheba tested Solomon's wisdom by challenging him to distinguish a flower from a wax facsimile thereof constructed in the royal workshop. As an aid to thought, suppose that she introduces these look-alikes to Solomon as Jacob and Esau—without, of course, telling him which is the artifact and which the flower. Then initially, before he determines, with the help of a bumblebee from the garden, that Jacob is the waxen artifact, Solomon finds it *believable* that Jacob should sprout new petals. Does he find this conceivable, though, in the sense relevant to possibility? Not if the stories about his wisdom are correct; he finds it undecidable on the available evidence. "If I assume that Jacob is a flower," Solomon might reflect, "then I *can* conceive it sprouting new petals; and if I assume that it is an artifact, then this becomes inconceivable for me. As it is, though, the petal hypothesis is neither conceivable nor inconceivable." Another story has Solomon ruling on a maternity case: is Mary, or Martha, the mother of this baby? Eventually he resolves the issue in Mary's favor, by offering to saw the baby in half. But initially, when Solomon found it believable both that Mary was the mother and that Martha was, did it appear to him that the baby's ancestry was metaphysically *contingent*? Only if such an appearance were compulsory could one maintain that believability entailed conceivability.

Two senses of "conceivable" have been distinguished: the believability sense (call it *conceivability_b*) and the philosopher's sense, the one that involves the appearance of possibility. Where the objector goes wrong is in failing to appreciate this distinction. Having uncovered a confusion about "could" in the argument from *conceivability_b* to possibility, he falls into a confusion of his own when he offers this as a refutation of *conceivability* arguments.

V. Some Circularity Objections

Suppose that we are careful to keep believability and conceivability apart, and that we conclude to *p*'s possibility only when *p* is conceivable. Even this would be bad procedure, if it could be shown that

conceivability is a guide to possibility *only as constrained by prior modal information tantamount to the information that p is possible.*

This is roughly what the circularity objection alleges. Because the objection is easily misunderstood, let me consider some things it had better *not* be saying before working up to what I think it is saying.

Even the staunchest defender of Hume's maxim would not insist that the conceivable was *always* possible, or that *p*'s conceivability *proved* its possi-

bility. Everyone is well aware of cases where impossible propositions have been found conceivable notwithstanding. The position to be defended, then, is only the following: that what is conceivable is typically possible, and that *p*'s conceivability justifies one in believing that possibly *p*.³⁵ Objection (A) does little more than reiterate these concessions in an accusing tone:

- (A) Since your argument is by admission fallible, you yourself recognize that it *might* fail in any given case. Therefore you should refuse to draw the conclusion, until you get prior assurances that it won't fail in *this* case. And that means: prior assurances that *p* is possible. So the argument becomes circular.

What is unconvincing here is the move from "the conclusion might be false, compatibly with the truth of the premise" to "you should refuse to draw the conclusion until you're sure that it is *not* false." Arguments like this *usually* lead from truth to truth, so unless there is reason to think that truth is *not* preserved, it makes sense to suppose that it is.

Do conceivability arguments have a deeper problem than ordinary fallibility? Maybe there is something special about their failures. If we think of an argument's premises as stating the *evidence* for its conclusion, it is an initially unsettling fact about conceivability arguments that when they fail, the evidence's very *existence* can be due to the conceiver's ignorance of the fact that her conclusion was false. So, Aristotle might not have been able to conceive matter as indefinitely divisible, if he had known that it could be divided only so far; "contingent identity" theorists like J. J. C. Smart might not have found mental and physical phenomena conceivable as distinct if they had realized that they were identical as a matter of necessity; and so on. For evidence to be in this sense *fragile* is hardly the usual thing. When Russell's chicken, for example, concludes from having been fed for months that he will be fed tomorrow, his evidence would *still* have existed even had he known his true fate. All the more striking, then, that when I conceive something in fact impossible, if I had appreciated its impossibility then the misleading evidence might not have been:

- (B) For all you know, you would not have found *p* conceivable if you had been better informed, specifically, if you had known that *p* was impossible. But evidence that might, for all you know, be dependent on ignorance is inherently untrustworthy. To be sure that your evidence is *not* thus dependent, you need to know that *p* is possible. But then your argument becomes circular: you must

³⁵ Further only *prima facie*, or *defeasible*, justification is claimed. Again, everyone knows of cases where additional evidence turns up that convinces us, or ought to, that *p* was not possible after all.

already know that *p* is possible, before you can conclude that it is from your ability to conceive it.

Now, it is a difficult question how fragile conceivability evidence really is. Whether foreknowledge of *p*'s impossibility would have prevented me from conceiving it seems to depend on how fully I grasp the *reasons* why *p* is impossible, and how revealing those reasons are. But let's assume, for argument's sake, that *whenever* I find an impossibility conceivable, I would not have done so, had I but realized the proposition's impossibility. The problem is to see why this should reduce my confidence that *this* conceivable proposition is possible. After all, I draw the modal conclusion because I take it that given my evidence, it's probably true. And how is that probability affected, if I agree that in those occasional cases where my conclusion is false, my evidence would not have existed if I'd somehow fastened on the truth beforehand? Such a circumstance makes my errors more *embarrassing*, perhaps, but it doesn't seem to make them any more *common*.³⁶

Some of the propositions I find conceivable are (I suppose) impossible, though of course I don't usually realize this in particular cases. Objection (B) tried to find a problem in the fact that my not realizing it is a *necessary* condition of my finding them conceivable. Maybe this gets things backwards, however. Maybe the problem is that my ignorance of these propositions' impossibility would *sufficiently* explain my ability to conceive them:

- (C) How can you infer to *p*'s possibility before you have ruled out alternative explanations of its conceivability? Since for *p* to be unbeknownst to you impossible would sufficiently account for your ability to conceive it, this is one of the alternative explanations you need to rule out. To rule it out, though, you need to know that *p* is possible, thus rendering the argument circular.

What is true in the objection is that when you base a claim on such and such evidence, the claim can be challenged by pointing to alternative explanations of the evidence which you are unable to exclude. They may have *looked* like ducks in the pond, but if there are known to be convincing decoy ducks about, you cannot *assume* that they were ducks unless you have something to say against the decoy hypothesis. There are limits, though. You are *not* required to rule out the alternative "explanation" that although they for some reason looked like ducks, in fact they were not, that is, that your evidence was somehow misleading. For one thing, this can hardly be considered an ex-

³⁶ Note that a *certain* degree of fragility is only to be expected with arguments of the *it appears that p / therefore p* variety. For instance, the dishes displayed outside some Japanese restaurants stop *looking* like food when you are told that they're plastic models. So it is not just conceivability appearances that sit uneasily with a full and proper appreciation of their deceptiveness.

planation of your evidence at all; for another, it is so far just allegation without the slightest reason to believe it. But how is objection (C) any better? The suggestion is that *perhaps* I had it appear to me that *p* was possible only because I somehow missed the fact that *p* was not possible. In short: perhaps my evidence is misleading. Perhaps it is, but don't I need a reason to think so before taking the idea seriously?

VI. The Circularity Objection

Actually, the last two objections were bound to fail. For notice a feature they have in common: they propose accounts of such conceivability errors *as in fact occur* but without addressing the issue of whether their occurrence is at all to be *expected*. When you *do* conceive an impossibility, they say, a necessary and/or sufficient condition for this is that you did not realize that it was impossible. But this is compatible with your conceiving impossibilities rarely or never. To make the case that you conceive them often, the premise the objector needs is not that ignorance of impossibility is all it takes to *explain* a conceivability error, assuming it made, but that such ignorance is all it takes to *make* one. This stronger premise can be motivated by looking at a second alleged fallacy in Descartes's argument for dualism—this one rather more interesting than the last.³⁷

From his conceivability as existing without a body, Descartes concludes that disembodied existence is possible for him. The fallacy is said to lie in the fact that he simply takes it for granted that he has no essential properties beyond those that he knows of.

Objections like this were put to Descartes repeatedly, most notably by Arnauld in the "Fourth Meditation." Arnauld's view is that

if the major premise of this syllogism [that the conceivability of *x* without *y* shows the possibility of *x* without *y*] is to be true, it must be taken to apply not to *any* kind of knowledge of a thing...; it must apply solely to knowledge which is adequate.³⁸

By *adequate* knowledge of a thing, Arnauld means knowledge of all of its essential properties. Although what is possible for Descartes depends on his essence in its entirety, what he can conceive of himself is constrained by just that portion of his essence that he knows of. Unless his self-knowledge is adequate, then, his capacity for incorporeal existence might, for all the thought experiment tells him, be obstructed by unappreciated necessary connections. Here is Shoemaker in the same spirit:

In the sense in which it is true that I can conceive myself existing in disembodied form, this comes to the fact that it is compatible with what I know about my essential nature...that I

³⁷ See also Yablo 1990.

³⁸ CSM II, p. 140, my interpolation and emphasis.

should exist in disembodied form. From this it does not follow that my essential nature is in fact such as to permit me to exist in disembodied form.³⁹

What concerns me here is not the viability of Descartes's specific argument, or the truth of its conclusion, but the strategy which Arnauld's (Shoemaker's) objection represents. To be consistent, Arnauld should hold that *no de re* conceivability intuitions are trustworthy, unless the ideas employed are certifiable in advance as adequate—as embracing every essential property of their objects. But then an enormous part of our modal thinking falls under suspicion.

No one would doubt of herself that (e.g.) she could have been born on a different day than actually, or lived in different places; and outside of philosophy, no one would question that we know such things. But how do we know them, if not by attempting to conceive ourselves with the relevant characteristics and finding that this presents no difficulties?

What gives this question its force is the specter of an Arnauldian skeptic who holds that, given the possible inadequacy of my self-knowledge, I am in no position to oppose even such patently absurd essentialist hypotheses as that I am essentially born on September 30, 1957. If I might, unbeknownst to myself, be essentially accompanied by my body, however clearly I seem able to conceive myself without it, why couldn't I also be essentially born on that day, however clearly I seem able to conceive myself born a day earlier or later? Equally open to question are conceivability intuitions about objects *other* than oneself, like my intuition that Humphrey could have been born on a different day or that the Eiffel Tower could be painted yellow; for here too the adequacy of my ideas has not been demonstrated. Really, the skeptic says, I have no basis to quarrel with *any* essentialist hypothesis about *any* object—even the superessentialist hypothesis that it could not have been different *in any way*—until I get assurances that none of the object's essential properties are hidden from me.⁴⁰

³⁹ Shoemaker 1984, p. 155.

⁴⁰ This brings out a seeming historical irony in Arnauld's position. Leibniz, in his correspondence with Arnauld, proposes that *none* of a thing's properties is accidental to it. Since Adam is such that Peter denied Christ some thousands of years after his death, this holds essentially of Adam, who would accordingly not have existed had Peter not gone on to be disloyal. Arnauld objects: "I find in myself the concept of an individual nature, since I find there the concept of myself. I have only to consult it, therefore, to know what is contained in this individual concept....I can think that I shall or shall not take a particular journey, while remaining very much assured that neither one nor the other will prevent my being myself" (Mason 1967, pp. 32–33). Within limits, we share Arnauld's assurance, but it is hard to see what entitles *him* to it. How does he know that his self-conception is adequate, i.e., that he is aware of *all* of his essential properties? To complete the irony, something uncomfortably like this Arnauldian point is put to Arnauld by Leibniz himself: "...although it is easy to judge that the number of feet in the diameter is not contained in the concept of a sphere in general, it is not so easy to judge with certainty...whether the journey which I plan to take is con-

At this point the restriction to *de re* propositions begins to seem artificial. If ignorance of an *individual's* essential properties can generate modal error, why not ignorance of a *property's* essential properties? Imagine that my grasp of a property *S* fails to reflect the fact that it is essentially uninstantiable (*S* might be the property of being sodium-free salt). Nothing to prevent me, then, from conceiving it as possible that *Ss* should exist: a *de dicto* conceivability error rather than a *de re* one. Likewise the *de dicto* impossibility that some *Qs* are *Rs* will be conceivable, if my understanding of *Q* omits its essential property of having no *Rs* in its extension. Probably there is *no* proposition for which a worry like this cannot be raised. In skeptical moods, Arnauld will always be able to point to a potential gap in my modal information that would enable me to find *p* conceivable despite its impossibility. This suggests one final generalization of his objection to Descartes:

- (D) If all it takes to find a proposition conceivable is to be unaware that it is impossible, then since impossibilities go unappreciated *all the time*, they are just as often conceivable. Before relying on conceivability evidence in any specific instance, then, you need a reason to think that in *this* case, *p's* conceivability signifies that it is possible rather than that, although it is impossible, you are unaware of this. That is, you need a reason to *deny* that

- (*) although you are unaware that *p* is impossible, *p* is impossible.

Because (*)'s first conjunct is true, and known to be—you *are* unaware that *p* is impossible—you can be reasonable in denying (*) only if you are in a position to deny its second conjunct. But its second conjunct is that *p* is impossible! So you must already know that *p* is possible before you can conclude that it is from its conceivability.

(D) is the strongest form I know of the circularity objection; my only doubts are about its opening sentence. That conceivability arguments are *fallible* is of course admitted. But all the Humean need claim is that they are reliable *enough* for me to say: I'm justified, because *probably*, if my evidence holds, then so does my conclusion. Have conceivability arguments really been shown to be *so* fallible that this can no longer be said?

Without claiming to know exactly how fallible that is, I use the word "often" so that if impossibilities are *often* conceivable, then conceivability

tained in the concept of me, otherwise it would be as easy to be a prophet as to be a geometer..." (*op. cit.*, p. 59).

evidence is not *per se* justifying. Here is the opening lemma spelled out more fully:

(E1) Almost always, when I am unaware that *p* is impossible, I find it conceivable.

(E2) Often, when *p* is impossible, I am unaware that it is impossible.

(E3) Often, when *p* is impossible, I find it conceivable.

The first sign of trouble is that (E)'s logical form

(F1) Almost all Bs are Cs.

(F2) Many As are Bs.

(F3) Many As are Cs.

is deductively invalid. From the premises we know only that there is a high concentration of Cs among Bs, and a significant concentration of Bs among As; what we don't know is whether these two concentrations *line up* to any significant extent. Thus it might be that although half of all As are Bs, only 1% of the Bs are As, and it is the *other* 99% of the Bs which make it the case that nearly all Bs are Cs. More generally, the Bs which are also As might form a small enough fraction of the total B-population to be subsumable under the allowable exceptions to the general rule that almost all Bs are Cs. This is illustrated by argument (G):

(G1) Almost all swimmers are fish. (say, 95%)

(G2) Many mammals are swimmers. (say, 50%)

(G3) Many mammals are fish. (0%)

The conclusion is false because the *mammalian* swimmers—the ABs—are one and all exceptions to the generalization that swimmers are usually fish—that almost all Bs are Cs.

As a rough but workable guide to when this kind of trouble arises, an argument of form (F) is acceptable just in case premise (F1) can be rewritten as

(F1*) Almost all Bs, *whether they are As or not*, are Cs

without loss of plausibility. Argument (G) is bad because when we rework the first premise as indicated, we get something false:

(G1*) Almost all swimmers, *whether mammals or not*, are fish.

Applying the rule to argument (E) yields

(E1*) Almost whenever I am unaware that *p* is impossible, *whether it is impossible or not*, I find it conceivable.

The question, then, is whether unawareness of impossibility is *uniformly* conducive to conceivability—whether the relation holds *regardless* of *p*'s modal status.

Take first propositions such that I am unaware that they are impossible *and they are possible*. Surely I *do* find a great many of *these* conceivable, including almost every possibility I claim knowledge of: that I could have been taller, for example, or a better dancer, or born on a different day.⁴¹ But the critical claim is that this generalizes to the *impossible* propositions:

(E1×) Almost always, when I am unaware that *p* is impossible, and it is impossible, I find it conceivable.

Because (E1×)'s antecedent says that I fail to appreciate the fact that *p* is impossible, this can be simplified to: unappreciated impossibilities are almost always conceivable.

Dialectically, at least, (E1×) is in a rather weak position. Remember that the objector is trying to convince someone not initially convinced of it that

(E3) Often, when *p* is impossible, I find it conceivable.

But anyone doubtful of (E3) will be doubly suspicious of (E1×), for understandable reasons. No one supposes that impossibilities *appreciated* as such are often conceivable; so to be doubtful that impossibilities are often conceivable is already to be doubtful that *unappreciated* impossibilities often are. And anyone doubtful that they are *often* conceivable will hardly be in a mood to concede (E1×)'s claim that they are *almost always* conceivable!

However the problem is more than dialectical. The objector makes a statistical hypothesis: that almost whenever you fail to appreciate a proposition's impossibility, you find it conceivable. Normally such hypotheses are advanced on the strength of confirming instances. Why not now? Part of the reason might be that hardly any exist. At least, almost every unappreciated impossibility one *knows* of—Goldbach's conjecture (or its denial), Jacob's sprouting new petals, Martha's maternity, etc.—is not conceivable but undecidable. Rather than enumerating cases, though, I issue a challenge: if we are as prone as the objector suggests to conceiving unappreciated impossibilities, I would like to know what some of them are.⁴²

⁴¹ Do I find conceivable *almost every* possibility such that I am not aware that it is impossible? Hardly—there are infinitely many unobvious arithmetical truths to the contrary—but let that pass.

⁴² Bearing in mind that *not* to find a proposition *inconceivable* is not yet to find it conceivable.

VII. Believability of Possibility

Where does the objector get his confidence that unappreciated impossibilities are almost always conceivable? Perhaps for him this is not a statistical hypothesis at all, but a consequence of what he means by conceivability.

To see what his definition might be, look again at Arnauld's complaint against Descartes: "how does it follow, from the fact that he is aware of nothing else belonging to his essence, that nothing else does in fact belong to it?" What is striking here is Arnauld's assumption that Descartes *thinks* it follows. After all, Descartes's premise is not that he is *unaware* that he is essentially embodied, it is that he can *conceive* himself in a disembodied condition. That Arnauld puts the one premise for the other suggests that at some level, he takes them to say the same: a conceivable proposition is just one not known to be impossible. Shoemaker is more straightforward:

in the sense in which it is true that I can conceive myself existing in disembodied form, this comes to the fact that it is compatible with what I know of my essential nature...that I should exist in disembodied form.

Apparently both authors equate conceivability, at least of the kind they find in Descartes, with what I will call *conceivability_{bp}*: the believability of *p* is possible.

Now, on *this* interpretation of conceivability, (E1 \times) looks awfully plausible. In fact it becomes something on the order of a conceptual truth, namely, that someone who doesn't realize that *p* is impossible will find its possibility believable. But if (E1 \times) is true on the new interpretation, then the critique of the last section no longer applies. What is my response to the circularity objection read in terms of conceivability_{bp}?

What response? I *share* the objector's doubts about conceivability_{bp} arguments. In fact let me throw in some additional doubts of my own. To find a proposition conceivable_{bp} is to find oneself unable to rule its possibility out. But you do not acquire justification for *believing* that something is possible simply through lacking justification for *denying* that it is. Otherwise, there could be no such thing as a person completely in the dark about *p*'s modal status; the less she knew *against p*'s possibility, the better her grounds would be for concluding that it *was* possible. (Recall that the argument from straight believability to possibility was criticized on similar grounds. If that argument was bad, the one from the believability of possibility is worse, for the new premise is strictly weaker than the old.)

So nothing as complicated as the circularity objection is needed to see that a proposition's possibility is not inferable from its conceivability_{bp}. But the objection's real problem is rather this: it *makes no difference* to Hume's maxim whether the inference goes through, for conceivability_{bp} fails the modal appearance test on both counts. Thus suppose that I have *no idea*

whether p is possible (p might be Goldbach's conjecture). Then I find p conceivable_{ep}—it is possible for all I know—but I have no inclination whatever to think it possible, nor have I misrepresented anything should it turn out not to be. In the end, then, the seemingly deeper circularity objection comes down to the same sort of misunderstanding as its predecessor: except that where the one mistook conceivability for the believability of truth, the other mistakes it for the believability of possibility.⁴³

VIII. The *A Posteriori* Objection

Up to now we have been looking at *traditional* criticisms of Hume's maxim. But some may feel that the really decisive difficulty came to light only recently, with the discovery by Kripke and Putnam of *a posteriori* necessary truths: that cats are animals, that Hesperus is identical to Phosphorus, and so on.⁴⁴ This would be strange if true, since for their own part these authors use conceivability methods all the time. But that is a separate issue; what is the problem that *a posteriori* necessary truths can *seem* to raise for the conceivability maxim?

Take any *a posteriori* necessity and negate it; the result is a necessary falsehood whose falsity is knowable only through experience, for instance, that cats aren't animals, or that water is distinct from H₂O. But, if it takes experience to show that these propositions are false, there ought to be *alternative* courses of experience that would have revealed them as true:

we can perfectly well imagine having experiences that would convince us (and that would make it rational to believe) that water is not H₂O. In that sense, it is conceivable that water isn't H₂O.⁴⁵

Putnam's conclusion is only that conceivability is no *proof* of possibility, but there is a more damaging result in prospect:

(G1) Whenever p is *a posteriori* false, I find it conceivable whether it is possible or not.

(G2) Often, *a posteriori* falsehoods are impossible.

(G3) So *a posteriori* falsehoods are often found conceivable despite their impossibility.

⁴³ This is not to say that Descartes's argument goes through. Perhaps Shoemaker is right that it is only in the believability-of-possibility sense that Descartes can conceive himself as disembodied. (Yet I assume that Descartes, for his part, would claim conceivability in a stronger sense; and so far we have no reason to doubt him.)

⁴⁴ See, for one, Teller 1984. By an *a posteriori* necessary truth I mean a necessarily true proposition whose truth is knowable *a posteriori* but not *a priori*; an *a posteriori* impossibility is the denial of an *a posteriori* necessary truth, in other words a metaphysical impossibility whose falsity is knowable only *a posteriori*.

⁴⁵ Putnam 1975, p. 233. For discussion purposes, I assume that water is necessarily H₂O.

This objection doesn't purport to embarrass *all* conceivability arguments, notice, only those where the conceived proposition is *a posteriori* false. But that is bad enough. For example, I should not argue from the conceivability of my sleeping late this morning, to the conclusion that this could really have happened. Even if it was *not* possible for me to sleep late, still I was going to find it conceivable that I should do just that.

IX. Epistemic Possibility

To conceive a proposition, in Putnam's sense, is to imagine acquiring evidence that justifies you in believing it: call this *conceivability_{ijb}*. But the definition is silent on a crucial point.

Distinguish three subtly different ways in which the thought experiment might go. Either the evidence is imagined to be disclosive of how things in the imagined situation really are; or it is imagined as for all its persuasiveness misleading; or whether the evidence is misleading is left unspecified. Speaking for myself, I can imagine being rationally persuaded of almost anything, provided I am allowed to imagine that the thing I am persuaded of is true, false, or of unspecified truth value, as I please.⁴⁶ To imagine a situation in which *p* is false, though, or one leaving *p*'s truth value unspecified, is *not* a way of having it appear to me that *p* could have been *true*. So the only relevant case, the only one where I am in danger of conceiving an impossibility, is the one where I imagine myself believing *p* justifiably *and truly*. That understood, justification becomes a side issue. For if the belief is imagined as *true*, then whether it is imagined as justified or not, my evidence for *p*'s possibility would seem to be exactly the same. (How could the imaginability of my *knowing* that *p* be better evidence of possibility than the imaginability of my truly believing it?)

Based on this reasoning, suppose we define *conceivability_{itb}* as the imaginability of veridically or truly believing that *p*. But, granted that this is different from *conceivability_{ijb}*, aren't *a posteriori* impossibilities *also* conceivable in the new sense? Can't I imagine truly believing that cats are robots, that Hesperus is distinct from Phosphorus, and so on?

Lurking just in the background here is a popular misunderstanding of Kripke's famous distinction between epistemic and metaphysical possibility. First it is emphasised that for Hesperus to have been other than Phosphorus is metaphysically impossible; *it could not have been* that Hesperus was not Phosphorus. Then it is explained that their nonidentity is nevertheless *epis-*

⁴⁶ Thus I can imagine some leading number theorist announcing an error in Euclid's proof from which it emerges that there is a largest prime number after all; the error takes years of training to understand but the authorities are convinced, and I, naturally, defer to their superior knowledge. Although my *imagined* self is convinced, my *actual* self is not; I find a largest prime unimaginable and so I suppose the imagined authorities to be mistaken.

temically possible, since it *could have turned out* that they were not the same.

All of this is correct but the last step: the explanation of what epistemic possibility consists in. 'It could have turned out that *p*' claims, I assume, either the possibility, or the imaginability, of our coming to believe that *p* and believe it truly. On the *first* reading, as Kripke says, "*it could have turned out that p entails that p could have been the case.*"⁴⁷ Since it could not have *been* the case that Hesperus and Phosphorus were distinct, they could not have *turned out* to be distinct. But, and this is the point, the explanation in terms of imaginability fares no better. To imagine myself truly believing that Hesperus and Phosphorus were distinct, I would have to imagine them *being* distinct; and that I cannot do, no more than I can imagine Venus's being distinct from Venus.⁴⁸

Now it is a given that all of the usual *a posteriori* impossibilities⁴⁹ are to come out epistemically possible; this is the result for which Kripke introduced the notion. Since *not* all of these *a posteriori* impossibilities are conceivable_{iib}—*Hesperus ≠ Phosphorus* was our counterexample—conceivability_{iib} cannot be what Kripke intends by "epistemic possibility." For much the same reason, though, conceivability_{iib} is not a good reading either of "conceivable" as it occurs in the *a posteriority* objection. Unless we find *a posteriori* impossibilities "conceivable," the objection proceeds from a false premise; and to repeat, we do *not* seem to find them conceivable_{iib}.

Still it is hard to shake the feeling that there is *some* worthwhile sense in which we *can* imagine truly believing that Hesperus is not Phosphorus, that cats aren't animals, and so on. Since that might be the sense the *a posteriority* objection is looking for, let us consider the matter one more time. What is it to imagine yourself truly believing something? To believe truly is to believe a truth, so you imagine a situation in which you believe some true proposition. On reflection, though, it is not completely obvious how this proposition is to be identified. Is it the proposition that your *hypothetical* self entertains when it inwardly pronounces, say, '*water ≠ H₂O*,' or the one that your *actual* self entertains? For these can be different.

Recall that the paper in which Putnam calls '*water ≠ H₂O*' a conceivable impossibility contains in addition a story about how propositional content is fixed. Which proposition I believe, Putnam says, is a function not only of what goes on "in my head"—my narrow psychological state—but also of extrinsic contextual factors, including, for instance, facts about my causal inter-

⁴⁷ Kripke 1980, pp. 141–42.

⁴⁸ "But we *could* imagine veridically believing them to be distinct, back when we thought they *were* distinct." True but irrelevant; it remains that *Hesperus ≠ Phosphorus* is *now* epistemically possible, but *not* now conceivable_{iib}.

⁴⁹ *Water ≠ H₂O*, gold is a compound, cats are robots, this lectern was originally made of ice, and so on.

actions with the larger world. Thus the narrow psychological state, internal mental act, or what have you, constitutes only my subjective *contribution* to propositional content.⁵⁰

How to fit beliefs themselves into the picture is a further question, and a disputed one. Some would individuate beliefs so that as long as the subjective contribution holds steady, the belief does too; variation in context affects not the belief *per se* but only the proposition believed. Others think of beliefs as having their propositional contents essentially: if I had believed a different proposition, then let my subjective condition be as similar as you like, I would have had a different belief. Rather than taking sides in this debate, suppose we concede the term “belief” to the second camp, and use “thought” to stand for the subjective contribution only. Thus my *thought* will be the internal state or act that determines, in context, which proposition I believe—what I will call the proposition *expressed* by the thought in that context. For instance, the thought which in the existing context expresses the proposition that Hesperus \neq Phosphorus, *would* have expressed a different proposition, a proposition with the truth conditions that Venus \neq Mars, if Mars rather than Venus had been responsible for the appearances by which the referent of ‘Phosphorus’ is canonically identified.

All of which brings us back to the original question: in imagining, or seeming to imagine, myself truly believing an *a posteriori* impossibility *p*, do I imagine myself believing the proposition that my *p*-thought *actually* expresses? or believing *some other* proposition, the one that my *p*-thought *would* have expressed had the imagined situation obtained?⁵¹

Start with the first option: imagining myself believing the proposition that my *p*-thought *actually* expresses. Since the proposition actually expressed by my *p*-thought is the proposition that *p*, this is just conceivability_{itb} again. What about the second option? Well, I can imagine believing *something* true with my Hesperus \neq Phosphorus-thought, for as I said, I can imagine it expressing a proposition with the truth-conditions that Venus \neq Mars. Since I *cannot* imagine myself truly believing that Hesperus \neq Phosphorus, we have uncovered a new kind of conceivability: *p* is *conceivable_{ep}* if one can imagine, not truly believing that *p* (that very proposition!!), but believing *something* true with one’s actual *p*-thought.⁵²

⁵⁰ See, for example, Dennett 1982 and White 1982.

⁵¹ Depending on one’s theory of propositions, the same proposition *p* could be expressible, in the same world and context, by distinct thoughts *t* and *t’* (so, the thought that the Morning Star \neq the Evening Star might express the same proposition as the thought that Venus \neq Venus). But then if someone thinks both *t* and *t’* on a given occasion, the phrase “her *p*-thought on that occasion” will be ambiguous between *t* and *t’*. I will not bother about this problem except to say that it vanishes if we treat epistemic possibility as a property directly of thoughts.

⁵² The subscript “ep” is for *epistemic possibility*. Some will regard the analysis as too weak, others as too strong.

How does the *a posteriori* argument look in light of these distinctions, in particular its leading premise that all *a posteriori* falsehoods are conceivable? Read in terms of conceivability_{ib} or conceivability_{ep}, the premise is not unreasonable. For an *a posteriori* falsehood to be conceivable in *these* senses therefore says little for its possibility. Remember, though, that Hume's maxim claims evidential import only for the kind of conceivability that *portrays p* as possible. And the kinds just mentioned do not: the appearances they convey are rather that you could have been justified in believing that *p*, and that you could have believed some truth or other via the thought you actually use to believe that *p*.⁵³ That leaves conceivability_{ib}. This *does* seem to involve the appearance of possibility, so Hume has some explaining to do if for all *a posteriori* falsehoods *p*, one can imagine truly believing that *p*. But this has not been argued, and as regards *a posteriori* impossibilities I doubt there are many who would even defend it. What we *can* do is imagine believing them justifiably, and believing related propositions truly; what we cannot do is imagine believing them, truly.

X. What Conceivability Is

Before attempting a positive account of conceivability, let me say something to lower expectations about what such an account should involve. Almost

Too weak: "What I find epistemically possible ought to be constrained by my immediate evidential situation. For instance, if I know my visual field to be wholly red, then it should not be epistemically possible that it is wholly green. Yet this is conceivable_{ep}; I can imagine believing something true with the thought that my visual field is wholly green, for I can imagine its *being* wholly green." To accommodate this intuition we might try the following. Define a thought as *cartesian* if it constitutes certain knowledge of the proposition it expresses, and it could not have expressed any other proposition; and let *c* be the conjunction of all propositions one thinks by way of cartesian thoughts. Then *p* is conceivable_{epc} if the conjunction of *p* with *c* is conceivable_{ep}.

Too strong: "Epistemic possibility ought to be a *weaker* notion than conceivability. Roughly it should be conceivability unconstrained by empirical beliefs. But some conceivable propositions are not conceivable_{ep}, for instance, the proposition that there are no thoughts." To accommodate *this* intuition, we need to arrange it so that thoughts continue to express propositions even in worlds where they do not exist. Say that the proposition a thought expresses in such a world is the one it expresses in the most natural expansion thereof to a world in which the thought *does* exist. Then *p* is conceivable_{epw} if one can imagine a world that verifies the proposition that one's *p*-thought expresses therein.

Kripke offers no explicit definition of epistemic possibility but his idea is that "under appropriate qualitatively identical evidential situations, an appropriate corresponding qualitative statement might have been [true]" (*op. cit.*, 142). This goes over into conceivability_{epc} if by "qualitatively identical evidential situations" we understand situations satisfying the conjunction of all propositions one thinks by way of cartesian thoughts; and by a "corresponding qualitative statement" to *p* we understand a proposition *p** such that *p** is true at a world *w* iff one's *p*-thought expresses a truth there.

⁵³ And these things presumably are possible when *p* is a *a posteriori* false.

never in philosophy are we able to analyze an intentional notion outright, in genuinely independent terms: so that a novice could learn, say, what memory and perception *were* just by consulting their analyses. About all one can normally hope for is to locate the target phenomenon relative to salient alternatives, and to find the kind of internal structure in it that would explain some of its characteristic behavior. This at any rate is all I have hopes of doing for conceivability—and so much the better, in my view, if it can be done while remaining as neutral as possible on other issues. This section and the next propose an account that locates conceivability proper with respect to the various subscripted impostors; makes for a revealing contrast with inconceivability and undecidability; predicts that a conceived proposition will appear as possible; and does little else besides.

Here are the five main conceivability-notions that we have considered so far. Each should really be relativized to a person and an occasion, but we will be sloppy:

- p is *conceivable_b* iff
it is (not un)believable that p .
- p is *conceivable_{bp}* iff
it is (not un)believable that *possibly*, p
- p is *conceivable_{ib}* iff
one can imagine justifiably believing that p .
- p is *conceivable_{itb}* iff
one can imagine believing p truly.
- p is *conceivable_{ep}* iff
one can imagine believing something true with one's actual p -thought.

What I have been calling philosophical conceivability is none of these. Conceivability in the imaginability-of-true-belief sense comes closest, but has the following problem. I cannot imagine truly believing anything that conflicts with the hypothesis of my believing it: that I do not exist, for instance, or that no one has any beliefs. Yet many such propositions are philosophically conceivable, including the ones just mentioned.

From the way I have presented the problem you can guess its solution: I find p conceivable if I can imagine, not a situation *in* which I truly believe that p , but one *of* which I truly believe that p . This is the approach to be developed in what follows. And the obvious place to begin is with the nature of imagination.⁵⁴

⁵⁴ For a fuller discussion that supports on some points the approach taken here, see Walton 1990.

Imagining can be either *propositional*—imagining that there is a tiger behind the curtain—or *objectual*—imagining the tiger itself.⁵⁵ To be sure, in imagining the tiger, I imagine it as endowed with certain properties, such as sitting behind the curtain or preparing to leap; and I may also imagine *that* it has those properties. So objectual imagining has in some cases a propositional accompaniment. Still the two kinds of imagining are distinct, for only the second has alethic content—the kind that can be evaluated as true or false—and only the first has referential content—the kind that purports to depict an object.⁵⁶

Objectual imagining, I said, may be accompanied by propositional imagining. But it is the other direction that interests me more: propositional imagining as accompanied by, and proceeding by way of, objectual imagining. To imagine that there is a tiger behind the curtain, for instance, I imagine a tiger, and I imagine it as behind the curtain. Quite possibly though I imagine the tiger as possessed of various additional properties—facing in roughly a certain direction, having roughly a certain color, and so on—and I imagine besides the tiger various other objects—the curtain, the window, the floor between them—all arranged so as to verify my imagined proposition. In short I imagine a more or less determinate *situation* which I take to be one in which my proposition holds. This is a closer approximation to what I mean by finding *p* conceivable; but “more or less determinate situation” is not quite right.

When I imagine a tiger I imagine it as possessed of some determinate striping—what else?—but there need be no determinate striping such that I imagine my tiger as striped like *that*; the content of my imagining is satisfiable by *variously* striped tigers, but not by tigers of *no* determinate striping. Likewise for situations: even if there is much about my tiger-situation that I leave unspecified as irrelevant to the proposition at hand (e.g., the distance from the tiger’s nose to the curtain), still I think of these things as fully definite in the situation itself. Thus a situation in which the tiger stands at *no* particular distance from the curtain, supposing that one can imagine this at all, is not what I have in mind.

⁵⁵ Some philosophers use “imagine” so that imagining a thing is *imagining* it, that is, conjuring up an appropriate sensory presentation. I do *not* require a sensory-like image for imagining, and certainly not a distinct such image for distinct imaginings. (Compare Descartes on the unimaginability of chiliagons at CSM II, pp. 50, 69, 264.)

⁵⁶ “Can’t the content of *objectual* imagining be truth-evaluable as well, if what one imagines is a proposition?”

This shows the importance of distinguishing the object of an imaginative act from its content. In the case described, the *object* of my imagining is a proposition. But its *content* is no more a proposition than the content of my tiger-imagining is a tiger. Rather it is something more on the order of a concept, the concept of being the proposition that a tiger behind the curtain is about to leap. Concepts being referential rather than truth-bearing, the criterion gives the right result.

By a *determinate* object, I mean one that possesses for each of its determinable properties an underlying determinate (it is not *merely* triangular, but in addition scalene, isosceles, or equilateral).⁵⁷ To imagine an object *as* determinate is to imagine it as possessing the higher-order property stated, that of possessing a determinate property for each of its determinables. There is a world of difference, then, between *imagining an object as determinate*—as possessing determinates for each of its determinables—and *determinately imagining it*—*specifying* in each case what the underlying determinate is. What I have been urging is that objectual imagining is determinate in the first sense but not the second. The one remaining question is whether the imagined object is itself indeterminate, as the phrase “more or less determinate situation” seems to suggest.

Suppose that it is, so that I imagine an indeterminate tiger rather than a determinate one. Then were a real, determinate, tiger to step out from behind the curtain, I ought to say that I had something more indeterminate in mind; whereas if an indeterminate tiger (!!) emerged, I ought to welcome it as just what I’d imagined. This of course gets things exactly backwards. Do I imagine a determinate tiger, then? Not if this means that I am *en rapport* with one of all possible tigers, striped in one of all possible ways, etc. But to repeat a point already made, it is one thing to imagine an object as being of such-and-such a type, another for there to be an object of that type such that one imagines *it*. Understood on the first and more natural model, “I imagine a determinate tiger” describes the case perfectly.

Why should it be different, if the imagined object is a situation rather than a tiger? What we are tempted to describe as imagining a *more or less* determinate situation, is better described as imagining a *fully* determinate situation whose determinate properties are left more or less unspecified.

When I imagine a situation, I imagine a completely determinate one. Is this the same as imagining a possible world? Unfortunately not quite. Possible worlds are situations complete in *every* respect: spatially, temporally, and ontologically, for instance. But from determinacy alone these other dimensions of completeness do not follow. I may indeed imagine my tiger-situation as *part* of a complete situation, including, besides the tiger and its immediate neighbors, everything that coexists with them all laid out in some nameless pattern. But although this larger reality is in a sense *acknowledged*—I think of my tiger-situation as embedded in it—the point of calling it *larger* is that I do not imagine the whole of it in imagining the tiger-situation *per se*.

That I imagine my tiger-situation as limited is slightly awkward for our plan of explaining conceivability as the imaginability of a situation in which

⁵⁷ Compare Locke’s account of the “general idea of the triangle” as triangular but “neither oblique nor rectangular, neither equilateral, equicrural, nor scalenon” (Locke 1959, book IV, chapter 7, section 9). Lockean general ideas, if they existed, would be indeterminate in the sense intended; likewise “arbitrary objects” as discussed in Fine 1983.

the conceived proposition is true. On the usual theory, propositions have truth values not in limited situations, but in the complete situations I have identified with possible worlds.⁵⁸ Luckily there is a way of correcting for this: As a rule, objectual imagining radically underdefines its object; so in principle it should be possible to imagine a *p*-verifying world while leaving matters visibly irrelevant to *p*'s truth value unspecified. Granted that this is not itself to imagine a (limited) *p*-verifying situation, the two imaginings are closely related and it would seem natural for them to occur together. To look at the matter from the other direction, even if imagining my tiger-situation is not the *same* as imagining its larger world, I may well imagine the larger world *in addition*. This latter imagining is of course hopelessly unforthcoming about events outside the tiger's immediate neighborhood, but so it would be if its mission was to arrange for the truth of a proposition indifferent to those events; and so it should be, if it is to go proxy for imagining a situation in which those events have no part. I propose that the work that might have been done by the imagining of *situations* in our analysis, can be done instead by the imagining of *worlds* understood mainly as containing those situations.

Now the pieces begin to fall together. Conceiving that *p* is a way of imagining that *p*; it is imagining that *p* by imagining a world of which *p* is held to be a true description. Thus *p* is conceivable for me if

(CON) I can imagine a world that I take to verify *p*.⁵⁹

Inconceivability is explained along similar lines:

(INC) I cannot imagine any world that I don't take to falsify *p*.⁶⁰

⁵⁸ Sense can also be made of truth-in-a-limited-situation, but it would be distracting to try to harmonize the two approaches here.

⁵⁹ It would be closer to ordinary language to distinguish 'I conceive that *p*,' '*p* is conceivable for me,' and 'I find *p* conceivable' as follows: (a) 'I conceive that *p*' iff I imagine a world which I take to verify *p*; (b) '*p* is conceivable for me' iff I *can* conceive that *p*; and (c) 'I find *p* conceivable' iff I find that I can conceive that *p*, presumably, by attempting to conceive it and finding that I succeed. But although my usage in this paper is *roughly* in accord with (a) and (b), to reduce clutter I have used 'I find *p* conceivable' and 'I conceive that *p*' more or less interchangeably. (Compare: 'I find it desirable' 'regrettable/acceptable...that *p*' is sometimes just a lengthier way of saying that I desire/regret/accept...that *p*.)

⁶⁰ *Objection:* Suppose *p* is the proposition that Socrates is a rain cloud. Then *p* is inconceivable to me; but I *can* imagine worlds that I don't take to falsify *p*, for I can imagine worlds in which Socrates doesn't exist. *Reply:* "Falsify" in (INC) is short for "fail to verify." For any world you can imagine, you take that world not to verify the proposition that Socrates is a rain cloud; hence you take it to "falsify" that proposition in the sense intended. To stress, on the intended reading (INC) is equivalent to the following: for every world I can imagine, I take that world not to verify *p*. (Brevity is not the only reason for using "falsify" rather than "fail to verify." The other reason is to discourage

Obvious as this account may seem, it leads in interesting directions; and as it is, it fares better than any other account I know with the modal appearance test.

Tigers with round-square striping are not imaginable; neither can we imagine tigers that lick all and only tigers that do not lick themselves, or tigers with more salt in their stomachs than sodium chloride, or indeed any tigers that do not strike us as capable of existing. Assuming that this is no coincidence, two explanations suggest themselves:

- (1) one cannot imagine an *X* unless it *already* appears to one that an *X* could exist; and
- (2) to imagine an *X* is *thereby* to enjoy the appearance that an *X* could exist.

Which of these is more plausible? If (1) were correct, then we could never *arrive* at the view that *Xs* are possible by succeeding in imagining one. Surely though this is the *usual* way of coming to regard *Xs* as possible. For instance, it is only by learning how to imagine such things that we admit the possibility of, say, justified true beliefs that do not rise to the level of knowledge, or physical duplicates of ourselves that mean different things by their words. This shows that it cannot be a prerequisite of imagining an *X* to be under the prior impression that *Xs* can exist. Which leaves (2) as the likelier explanation: it *comes* to me that *Xs* are possible in the act of imagining one.⁶¹

Assuming that objectual imagining works the way (2) says, it is no mystery why conceiving, in the sense of (CON), involves the appearance of possibility. By (2), when I imagine a world of such and such a type, it appears to me that a world of that type could really have existed. But when I take it to verify *p*, I take it that if a world like that *had* existed, then *p* would have been the case. So, when I imagine a world which I take to verify *p*—and this is what it *is* to conceive that *p* on the proposed account—I have it appear to me that *p* is possible.

XI. Undecidability

Part of the appeal of (CON) and (INC) taken together is that they leave room for a *third* conceivability-status, such as undecidability was supposed to be. At least there is no obvious contradiction between

confusion with the much weaker condition that: for every world I can imagine, I do not take that world to verify *p*. This latter condition defines *nonconceivability*.)

⁶¹ Or, if this seems debatable, I hereby *stipulate* that “imagining an *X*” will denote type-(2) imagining.

- ($\overline{\text{CON}}$) I cannot imagine a world that I take to verify p , and
 ($\overline{\text{INC}}$) I can imagine worlds that I don't take to falsify p ;

and since these are the denials of (CON) and (INC), their conjunction defines undecidability. But although ($\overline{\text{CON}}$) and ($\overline{\text{INC}}$) are formally consistent, someone might still wonder how both could be true at the same time. For this would require that in attempting to conceive that p , I find myself imagining worlds such that it is *obscure* to me whether they verify p or falsify it. And do cases like this actually arise?

According to (CON), the task of conceiving p divides into two sub-tasks: imagining a possible world, and satisfying oneself that p is true in it. Often the world can be *stipulated* to be one in which p is true, as for example when Kripke stipulates that the man imagined to be President is our own Hubert Humphrey; then the verification task is trivial. But for some values of p , worlds in which p is clearly true are not clearly imaginable, or, what comes to the same, in clearly imaginable worlds p 's truth-value seems somehow uncertain. So, given his problems imagining a world in which Jacob sprouts new petals, Solomon may seek firmer ground in the hypothesis of a world where Jacob acquires petal-like appendages—whether these are petals is left obscure in deference to the possibility that Jacob is an artifact. Because he can imagine no world that he is ready to count as one in which Jacob sprouts new petals, the Jacob-proposition is not conceivable for him; but neither is it inconceivable, for he *can* imagine worlds which he is unready to describe as ones in which the proposition is false.

Another proposition I have called undecidable is not-GC, the denial of Goldbach's conjecture. Many philosophers have suggested that not-GC is rather conceivable. Michael Hooker, for instance, writes that one can

imagine the discovery by computer of a counterexample to the conjecture, the attendant discussion of it, the subsequent revision of philosophical examples, etc.⁶²

To explain where I think this goes wrong, let me describe some scenarios I clearly *can* imagine and then show how imagining *these* falls short of imagining that not-GC. For instance, I find it easy to imagine a computer printing out some unspecified even number n , and this being hailed on all sides as an authentic counterexample. Why wouldn't this be a case of imagining that not-GC? Because it suffices for the veridicality of *this* imagining for the following to be possible: GC has no counterexamples, but the computer produces a number n widely though erroneously hailed as a counterexample.

⁶² Hooker 1978, p. 178.

Thus the truth of my imagining does *not* depend on there being a world in which not-GC, as it would if I had succeeded in imagining that not-GC.

Maybe I do better to imagine the computer producing something widely acknowledged as a *proof* that n is a counterexample. But again, the proof can help me to enjoy the appearance that possibly not-GC only if it is imagined to be *correct*; and since it is inconceivable to me that addition facts should vary between possible worlds, my ability to imagine the proof as correct is limited by my confidence that some number is *in fact* unavailable as the sum of two primes. Alas, I have no idea whether such a number exists, and neither (I assume) does anyone else. How then can I treat the computer's output as a correct proof? Am I to imagine it set out in convincing detail? But if the detail is only *imagined* to be convincing, it does nothing to increase my *actual* confidence in the proof's correctness. Am I to imagine the proof set out in *actually* convincing detail? If I could, I would call a press conference to announce my refutation of Goldbach's conjecture! So no Hooker-type thought experiment that I'm aware of shows the conceivability of not-GC. What the thought experiments *do* suggest is that not-GC is not *inconceivable*; accordingly it is undecidable.

XII. Modal Error

Ordinarily we treat perceptual appearances as *prima facie* accurate, and absent specific grounds for doubt we accept them as a basis for reasonable belief. What about conceivability appearances? Outside of philosophy, at least, they are treated in a similar fashion. Suppose that you claim to be able to imagine a world in which Oxford University exists but Cambridge does not. Perhaps we can point to some complicating factor of a kind you had not considered, e.g., one was originally a college of the other, that takes our own modal intuitions in a different direction. But if nothing of the kind occurs to us, and if attempting the thought experiment ourselves we find no difficulty in it, we are in a poor position to dispute your claim. (Imagine your reaction if we said, "still, we wonder if it is really possible," though no further complication suggested itself.)

So common sense sees appearances of both kinds as *prima facie* accurate and *prima facie* justifying. About conceivability appearances philosophers have taken a different view, but for unconvincing reasons. Can we stop worrying, then, and modalize with a clear conscience?

What makes us hesitate is not that conceiving can sometimes lead astray, but that we have so little idea how this happens. Modal error is a fact of life, and although perceptual error is too, our firmer grip on its etiology allows us to feel less the helpless victim than in the modal case. Misperception is something that we know how to guard against, detect when it occurs, and explain away as arising out of determinate cognitive lapses. That there is noth-

ing remotely comparable for conceivability is a measure of our relative backwardness on the subject of modal error. Of course, the analogy with perception can be taken too far; a more realistic comparison might be with mathematics. Yet the system of checks and balances in mathematics is in its way most impressive of all and certainly well beyond anything encountered in the modal domain.

No wonder the advice to “trust your modal intuitions” sounds overeasiness. Until our imaginative excesses are brought under something *like* the epistemological control we have in other areas, we modalize with right, perhaps, but without conviction.

Whatever their other problems, our objections at least had *models* to offer of how modal intuition goes wrong. Probably the most familiar is the one associated with the circularity objection: because you didn’t appreciate *p*’s impossibility, there was nothing to prevent you from finding it conceivable. Even if this particular explanation disappoints, some *such* explanation is badly needed.

How *does* it happen that people find (what are in fact) impossible propositions conceivable? Maybe it looks like I’ve ruled modal error out altogether! Because what I’ve said is that when a proposition is unbeknownst to me impossible, it is not normally inconceivable for me but undecidable.—*Normally*, but not always. The ancient Greeks, believing that Hesperus and Phosphorus were different planets, might well have found it conceivable for the one to outlast the other. That was a mistake; Hesperus is identical to Phosphorus, so they could not have been different in any way. Or suppose that Oedipus, upset with Jocasta, finds himself imagining what life would have been like without her. Even if she had never existed, he decides, he could still have been king. Assuming with Kripke that ancestry is essential, he could not have been *anything* if she had never existed; so here is another example of modal intuition misfiring.

Sure as I might be, then, that my modal intuitions are *largely* reliable, in any particular case I have the following worry. Sometimes people have found impossibilities conceivable. Maybe I am making an analogous error when I imagine myself born on October 1, or six feet tall, or a Rosicrucian, and conclude that these things are possible for me.

XIII. Models of Modal Error⁶³

Is the analogy a good one, though? Remember that the ancients found it conceivable that Hesperus should outlast Phosphorus only because they took it that Hesperus and Phosphorus were distinct. What is the prior misapprehension that accounts for *my* erroneous intuition, as the ancients’ denial of Hesperus’s identity with Phosphorus accounts for theirs?

⁶³ This section and the next are based on Yablo 1990.

That the request for a backing misapprehension sounds so reasonable suggests the following model of modal error.⁶⁴ First, I find p conceivable, when as a matter of fact it is impossible. Second, that p is impossible emerges from the truth of some proposition q . Third, I do not realize this, believing instead that q is false, or else that it is false that if q , then p is impossible; and this is how I am able to conceive p despite its impossibility. Explicitly, there is a proposition q such that ...

- (a) q ;
- (b) if q , then $\Box \sim p$; and
- (c) that I find p conceivable is explained by my denial of (a) and/or my denial of (b).

(‘ $\Box s$ ’ means: necessarily, s .) So, the ancients conceived it as possible for Hesperus to outlast Phosphorus because they denied the truth that Hesperus is identical to Phosphorus. If some contemporary philosophers, aware of this identity, find themselves capable of the same conception, the probable explanation is that they deny that identicals are modally indiscernible, and more particularly that Hesperus’s identity with Phosphorus makes a difference in lifespan impossible. In our other example, Oedipus’s false belief that Jocasta is not his mother explains how he can conceive himself being king even if she had never lived. Should he persist in his error after his ancestry is revealed, this is because he denies that *if* Jocasta is his mother, then he could not have been king without her.

Whatever you find conceivable, you are *prima facie* entitled to regard as metaphysically possible. The question is whether this *prima facie* entitlement can be defeated along the lines just indicated. Of course, if someone can *prove* that the model applies, then since (a) and (b) entail that p is impossible, your conclusion is refuted. But to raise legitimate *doubts* about the conclusion, reason to think that the model *may* apply ought to be enough. Thus we call proposition q a *defeater* if there is a reasonable chance that (a), (b), and (c).^{65,66} The objector’s challenge, in any particular case, is to find a defeater q of the conceiver’s modal intuition.

⁶⁴ Note: I do not say that *all* modal errors are captured by the models to be given here, only that many are, and especially the type most often discussed in recent modal metaphysics (see also note 67 below).

⁶⁵ Although it would be more in accord with existing usage to let the defeater be the *conjunction* of (a), (b), and (c). See note 67.

⁶⁶ How do I test the credibility of the conditional claim (b) that if q , then p is impossible? With any other indicative conditional, I use the Ramsey test: I pretend that I am reliably informed of the antecedent, and then I consider, under that pretense, how plausible I find the consequent. The same method works here. Suppose I want to decide whether, if salt = sodium, it is impossible for the ocean to contain more sodium than salt. Pretending that salt = sodium, I find it inconceivable that the ocean should contain these in different amounts; abandoning the pretense, I endorse the conditional.

Someone might object as follows. To erroneously conceive p as possible, why should I have to go so far as to *deny* the proposition q given which p is impossible, or to *deny* the proposition that p is impossible if q is true? Isn't it enough if I am simply *unaware* that q , or *unaware* that if q is true, then p is impossible? Thus consider a second, less demanding, model of modal error: there is a proposition q such that

- (a) q ;
- (b) if q then $\Box \sim p$; and
- (c) that I find p conceivable is explained by my unawareness that (a), and/or by my unawareness that (b).

Arguably this unawareness model does do a certain justice to cases which the denial model leaves untouched. At one time, for example, I suppose I found it conceivable that there should be a town whose resident barber shaved all and only the town's non-self-shavers. However it was not because I *denied* that the scenario was implicitly contradictory that I found the town conceivable; it was because I was not *aware* of the contradiction. Or imagine that the medievals, rather than denying that dolphins were mammals, had no opinion on the matter; suppose if you like that the concept of a mammal was unknown to them. Mightn't they have conceived it as possible, erroneously mind you, for dolphins to be fish? If so, then this would be another example of a false intuition whose explanation lay not in the fact that something was denied, but in the fact that it was not believed.

As before, the objector's challenge is to identify a proposition q for which there is a reasonable chance that the model applies.⁶⁷ Nothing could be easier, you might think. Just let q be the proposition that p , the proposition conceived, is impossible. Then since the conceiver's intuition is still *sub judice*, there would seem to be a reasonable chance that (a) q , that (b) if q , then p is

⁶⁷ This is a good place to acknowledge that the models given here cannot claim to accommodate *all* defeaters. Suppose we distinguish *rebutting* defeaters, propositions s such that $(\text{con}(p) \ \& \ s)$ is a reason to think that p is impossible; *offsetting* defeaters, propositions s such that $(\text{con}(p) \ \& \ s)$ is not a reason to think p possible; and *undermining* defeaters, propositions s such that s is a reason to deny that $\text{con}(p)$ is a reason to think p possible. And suppose we refer to conjunctions of (a), (b), and (c) as *standard* defeaters. Then standard defeaters are rebutting and offsetting (in virtue of (a) and (b)) and also undermining (in virtue of all three conjuncts). But none of our three categories is exhausted by the standard defeaters. For instance, intuition recognizes offsetting and undermining defeaters that are not rebutting. *Some* such are obtainable by generalizing the models to allow standard defeaters of $\text{con}(p^*)$, where p^* is a fuller description than p of the imagined world as the conceiver understands it. But even this leaves no room for defeaters like the following: you conceived that p while under the influence of a mind-expanding drug; your modal intuitions are famously inaccurate; everyone but you finds p undecidable.

impossible (this is a tautology), and that (c) the conceiver's ignorance of (a) explains how she managed to conceive *p* as possible.

Yet I take it that it gives me *no* reason to mistrust my intuition that *p* is possible to be told that it might, for all I know, be due to ignorance of what might, for all I know, be the fact that *p* is *not* possible; for instance, that my ability to conceive myself with a different birthday might derive from my failure to appreciate the necessity of my actual birthday. At best the objector can argue that *if* I am necessarily born on September 30, then my failure to realize this may be relevant to my finding a later birthday conceivable. And this hardly constitutes an *objection*, no more than it is an objection to the accuracy of my impression that there are ducks around that if I am wrong, and they are decoys, then my ignorance of that fact might help to explain how I managed to take them for ducks.

Part of my point here is just that ignorance of the fact that *p* is impossible does not itself do much to explain why I would conceive it as possible. But that is not all. Even if a fuller explanation is provided, it carries little dialectical force if it depends on the prior concession that my intuition has a significant chance of being false. (With equal plausibility one could explain away my perceptual impression of ducks by saying that they were produced by decoy ducks, these being the usual explanation of erroneous duck-impressions.) Only if there is *independent* reason to suspect that my refusal of some relevant proposition really does put me out of touch with the facts, does that refusal call my intuition into question.

XIV. Modal Dialogue

To see how this works in practice, consider again my Cartesian intuition that I can exist in a purely mental condition. Someone might object that it is independently plausible that I am embodied, and that if so, I am embodied necessarily and so incapable of purely mental existence. About the second half of this, I have my doubts. Like most people, I take it for granted that I *am* embodied. Somehow, though, this does not seem to inhibit me from conceiving myself as disembodied. This intuition of being actually-but-not-necessarily embodied *prima facie* rationalizes my *rejection* of the conditional hypothesis stated; so I cannot regard that hypothesis as independently credible. Of course, the conditional hypothesis becomes virtually certain if we let *q* be the proposition that I am *necessarily* embodied. Now though it is *q* itself which wants for independent evidence.

Another candidate for the role of defeater is that I am the *same thing* as my body. But what does "same thing" mean here? If it means identical, then I doubt that the defeater is independently plausible. However *categorically* similar my body and I may be, this suggests at most that we are *coincident* (as a

statue might be coincident with the hunk of clay that makes it up).⁶⁸ Evidence that we were moreover identical would be evidence that we agreed on a wide range of hypothetical, and especially *modal*, properties. Yet this can only come from conceivability considerations, which seem in fact to argue the other way! If “same thing” is understood so as to require sharing of categorical properties only, then the problem is just relocated. For now I need a reason to think that if I am categorically similar to my body, then I cannot exist without it. And to insist that categorical similarity has this consequence seems to beg the question against the otherwise intuitive view that what I am is a *person*, whose categorical properties may be those of a certain body, but with modal properties all my own.

Obviously the debate could be taken a lot further. To mention just two of the more promising possibilities, someone might try to extract a defeater from Kripke’s claim that my biological origins are essential to me, or from some version of the mental/physical supervenience thesis. But already we have enough to see how modal dialogue typically proceeds on the picture I have in mind.⁶⁹

- *X* finds *p* conceivable and calls it possible;
- if *Y* chooses to challenge *X*’s intuition, she proposes a defeater *q* to explain how *X* was capable of it despite its falsity;
- if *X* is unable to accept this explanation, he takes issue either with *q* itself, or with *Y*’s claim that it casts doubt on his intuition’s accuracy.

What to say—what it means—when the dialogue breaks down is the topic of the next section.

XV. Factualism about Modality

To defeat a modal intuition, the objector tries to motivate on independent grounds the suspicion that it derives from some prior error or oversight. Yet if conceivers disagree on fundamental enough matters—color incompatibilities, say, or the modal properties of mathematical objects—it may be difficult for either to discern on the other’s part a prior lapse at all, still less one independently recognizable as such. This raises the specter of brute modal error

⁶⁸ On coincidence and the categorical/hypothetical distinction, see Yablo 1987. I assume for the sake of the objection that there are no temporal differences between myself and my body, for instance, that my body isn’t going to outlast me.

⁶⁹ For reasons explained in note 67, the framework cannot be regarded as fully general. For instance, it doesn’t cover the case where *Y* challenges *X*’s intuition on the basis that he was drugged, or that he has often been wrong before. But I believe that it covers *most* modal disputes of the kind that arise between basically competent conceivers.

and disagreement. Too much of *that*, someone might say, and we lose the right to speak of error and disagreement at all.

Supporting this accusation is a theory of what it is for the statements in a given region of discourse to be genuinely factual, viz. that "differences of opinion about such statements...will have to be traceable back to some breach of ideal rationality or material difference in the subjects' respective states of information."⁷⁰ Reason to think that there is *just no saying* how the opposition comes by its seemingly *equally* well-supported conclusions despite their falsity is "reason to think that the statements disagreed about are not objective, and so not apt to be substantially true or false."⁷¹

Roughly, then, the proposal is to define factual discourse by its intolerance of brute error and disagreement. There are stronger and weaker versions of this, of course, and much that could be debated in all of them, but it is hard not to feel some sympathy for the basic idea. Unless the positions one would like to call incorrect show *some* tendency to be reproachable on separate grounds, the faith that there is anything genuinely at issue can indeed become strained. The alternative is to insist on there being "facts of the matter" that only oneself and one's coreligionists are privy to—that others, through no fault of their own, get consistently wrong. And although facts like that may not be unintelligible, they do have something of a credibility problem. This is especially so when, as in the modal case, our best idea of the type of fact in question is that of an external constraint on the outcome of a certain type of investigation: in the modal case, investigation by imagination. For then our confidence that there are facts of that type in play will be limited by our confidence that an external constraint really operates; hence by our resources for explaining how, despite the constraint, we are able to arrive at opposing views.

So, our entitlement to modal factualism turns on the effectiveness of our strategies against conflicts, or seeming conflicts, of conceivability intuition.⁷² (Here and below I use "conceivability intuition" broadly, as covering conceivability and inconceivability intuitions both.) What are those strategies? From the discussion above we have the following:

- (1) try to show that there is no conflict of *conceivability* intuitions because what looked like *p*'s conceivability was really only its believability, or epistemic possibility, or...; or what looked like its

⁷⁰ Wright 1986, p. 198. This is what Wright used to call the "rational command" criterion and now calls "cognitive command."

⁷¹ Wright 1988, p. 39.

⁷² At least, a certain *degree* of factualism might be in order if the condition were met. In his 1988 and elsewhere, Wright sketches a system of increasingly ambitious factualisms, and offers criteria appropriate to each. Here I employ a variant of his weakest criterion. Whether modal discourse is factual in his more ambitious senses I do not discuss; Wright himself is skeptical.

inconceivability was really only its unbelievability, or epistemic impossibility, or...;

- (2) admit that there are conceivability intuitions on either side but try to show that they are not in *conflict* because what seemed to be the conceivability (inconceivability) of one proposition was really that of some closely related other;
- (3) admit that there is a conflict of conceivability intuitions but try to show that at least one of them has a defeater and is therefore open to doubt.⁷³

(1) was the strategy we used with Goldbach's conjecture, when we said that it was "conceivable" only in the believability or the believability-of-possibility sense. The supposed intuition that Hesperus might not have been Phosphorus can be met with (1)—you find their nonidentity not conceivable but epistemically possible—or, what comes to the same in this case, (2)—it is not their nonidentity that you find conceivable, but only that you should have thought *something* true with your *Hesperus*≠*Phosphorus*-thought.⁷⁴ Another, more mundane, version of strategy (2) is to say that because of unnoticed idiolectic differences, the disputants talk past each other. Thus if we seem to disagree on the conceivability of a wet mop that holds no water, a possible explanation is that owing to differences in our concepts of wetness, the proposition I find inconceivable is not the one you find conceivable. (Sadly it is all too easy to believe that much of the current controversy over conditions of personal identity and survival—are teletransportation, brain transplant, mitotic division, etc. survivable?—owes more to our meaning slightly different things by "person" and "survive" than to any real clash of modal intuition.)

When the dissolving strategies fail, our one remaining option is to explain the conflict as arising out of some antecedent error or omission on one side or the other. To the newly crowned Oedipus, it seemed possible that he should have been king even if Jocasta had never existed; but what would you expect of someone deceived about his ancestry? The reason why some can conceive a barber who shaves all and only the non-self-shavers, while others

⁷³ To apply this strategy on the conceivability side of the conflict, we use the (a)(b)(c) model as presented in the text; to apply it on the inconceivability side, we extend the (a)(b)(c) model to inconceivability intuitions in the obvious way. Suppose that historians discover that Cicero was in reality Tully's older brother (that *q*), but that unaware of this I continue to find it inconceivable that the one should have outlived the other (that *p*). My intuition is defeated because (a) *q* is true; (b) if *q* is true, then *p* is possible; and (c) I find *p* inconceivable only because I am under the misimpression that *q* is false.

⁷⁴ In the case of the ancients, who really *did* find it conceivable that Hesperus should have been distinct from Phosphorus, strategy (3) is used: they were capable of this conception only because they were empirically and/or philosophically misinformed.

find this inconceivable, is that the first group needs to learn more logic. And so on.

But I have been putting off the essential question: what if, after all the strategies have been tried to the best of current knowledge and ability, there remains a residue of so-far-irreducible disagreement? Well, the factualist can say, there is still such a thing as *committing* ourselves to applying them in ever more inventive ways until one finally succeeds, or, failing that, to devising new and better strategies in a similar spirit. Such a commitment could of course come to seem awfully lame, if the failures proved stubborn and the successes too minor to balance them off. But there is another scenario I like better.

How is it that substantive modal metaphysics, after years in the doldrums, has lately been making headway again? Part of the explanation might be that our methods of modal conflict management have been in a real sense improving. Already it takes an effort to recall the dispiriting conditions of, say, thirty years ago: the various half-related ideas jumbled unconsciously together under the headings of possibility and conceivability; how crude the controls were on propositional content; the anxiety about collateral information as a factor in imaginability. Especially one forgets how much easier it was then for the conversation to bog down at the first clash of modal intuition. The extent to which we have moved beyond this should not be exaggerated (more often than not we *still* bog down), but meanwhile it seems that modal dialectic has achieved an unaccustomed degree of clarity and system in a surprisingly short time. All of this has been a tremendous boost to the factualist's morale; sufficiently more of it and her commitment above might well be vindicated. "But what is the verdict? Can modal metaphysics be brought under the discipline characteristic of a fact-finding enterprise or can't it?" I have no answer but just a suggestion: we should try to impose that discipline in the hope that it might eventually take.

REFERENCES

- Bealer, G. 1987: "The Limits of Scientific Essentialism." *Philosophical Perspectives* 1.
- Blackburn, S. 1986. "Morals and Modals," in G. MacDonald and C. Wright, ed., *Fact, Science, and Morality*. Oxford: Oxford University Press.
- Bradley, F. H. 1969: *Appearance and Reality*. Oxford: Oxford University Press.
- Chisholm, R. M. 1957: *Perceiving: A Philosophical Study*. Ithaca: Cornell University Press.
- van Cleve, J. 1983: "Conceivability and the Cartesian Argument for Dualism." *Pacific Philosophical Quarterly* 64, pp. 35-45.

- Coppock, P. 1984: "Review of N. Salmon, *Reference and Essence*," *Journal of Philosophy* 81, pp. 261-70.
- J. Cottingham, R. Stoothoff, & D. Murdoch, ed. 1985: *The Philosophical Writings of Descartes* (I, II). Cambridge, England: Cambridge University Press. (= CSM)
- Dennett, D. 1982: "Beyond Belief," in A. Woodfield, ed., *Thought and Content*. Oxford: Clarendon Press.
- DeRose, K. 1991: "Epistemic Possibilities," *Philosophical Review* 100, pp. 581-605
- Dreyfus, H. ed. 1982: *Husserl, Intentionality and Cognitive Science*. Cambridge, Massachusetts: Bradford.
- Fine, K. 1983: "A Defense of Arbitrary Objects," in *Proceedings of the Aristotelian Society*, supp.vol. 17.
- Forbes, G. 1985: *The Metaphysics of Modality*. Oxford: Clarendon Press.
- Hooker, M. 1978: "Descartes's Denial of Mind-Body Identity," in Hooker, ed., *Descartes: Critical and Interpretive Essays*.
- Hume, D. 1968: *Treatise of Human Nature*. Oxford: Clarendon Press.
- Hume, D. 1963: *An Enquiry Concerning Human Understanding*. La Salle: Open Court Press.
- Kneale, W. 1949: *Probability and Induction*. Oxford: Clarendon Press.
- Kripke, S. 1980: *Naming and Necessity*. Cambridge, Massachusetts: Harvard University Press.
- Locke, J. 1959: *An Essay Concerning Human Understanding*. New York: Dover.
- Mason, H. T., ed. 1967: *The Leibniz-Arnault Correspondence*. Manchester: Manchester University Press.
- Mill, J. S. 1874: *A System of Logic*. New York: Harper & Brothers.
- Mill, J. S. 1868. *An Examination of Sir William Hamilton's Philosophy*. Boston: W. V. Spenser.
- Moore, G. E. 1966: "Certainty," in his *Philosophical Papers*. New York: Macmillan.
- Pap, A. 1958: *Semantics and Necessary Truth*. New Haven: Yale University Press.
- Putnam, H. 1990: "Is Water Necessarily H₂O?" in his *Realism With A Human Face*. Cambridge, Massachusetts: Harvard University Press.
- Putnam, H. 1975: "The Meaning of 'Meaning'," in his *Mind, Language, and Reality*. Cambridge, England: Cambridge University Press.
- Reid, T. 1969: *Essays on the Intellectual Powers of Man*. Cambridge, Massachusetts: MIT Press.
- Searle, J. 1983: *Intentionality*. Cambridge, England: Cambridge University Press.

- Sellars, W. 1963: "Phenomenalism," in his *Science, Perception, and Reality*. London: Routledge & Kegan Paul.
- Shoemaker, S. 1984: "Immortality and Dualism," in his *Identity, Cause, and Mind*. New York: Cambridge University Press.
- Sidelle, A. 1989: *Necessity, Essence, and Individuation*. Ithaca: Cornell University Press.
- Teller, P. 1984: "A Poor Man's Guide to Supervenience," in *Southern Journal of Philosophy*, supp. vol. 22.
- Walton, K. 1990: *Mimesis as Make-Believe*. Cambridge, Massachusetts: Harvard University Press.
- White, S. 1982: "Partial Character and the Language of Thought." *Pacific Philosophical Quarterly* 63, pp. 347–65.
- Wright, C. 1988: "Realism, Antirealism, Irrealism, Quasi-realism," in *Midwest Studies in Philosophy* XII.
- Wright, C. 1986: "Inventing Logical Necessity," in J. Butterfield, ed., *Language, Mind and Logic*. Cambridge, England: Cambridge University Press.
- Yablo, S. forthcoming: "Review of Sidelle, *Necessity, Essence and Individuation*." *Philosophical Review*.
- Yablo, S. 1992: "Mental Causation." *Philosophical Review* 101, pp. 245–80.
- Yablo, S. 1990: "The Real Distinction Between Mind and Body," in *Canadian Journal of Philosophy*, supp. vol. 16.
- Yablo, S. 1987: "Identity, Essence, and Indiscernibility." *Journal of Philosophy* 84, pp. 293–314.