



Being Necessary: Themes of Ontology and Modality from the Work of Bob Hale

Ivette Fred-Rivera (ed.), Jessica Leech (ed.)

<https://doi.org/10.1093/oso/9780198792161.001.0001>

Published: 27 September 2018

Online ISBN: 9780191866876

Print ISBN: 9780198792161

Search in this book

CHAPTER

13 Counter-Conceivability Again

Crispin Wright

<https://doi.org/10.1093/oso/9780198792161.003.0013> Pages 266–282

Published: September 2018

Abstract

This chapter is concerned with the epistemology of metaphysical possibility implicit in the famous argument against physicalism about the mental outlined in the third lecture of Saul Kripke's *Naming and Necessity*. Kripke's argument presupposes that conceivability remains the best possible indicator of possibility, even where it is metaphysical possibilities, rather than conceptual possibilities, that are concerned. The chapter argues that this principle is good only when the concepts which frame the relevant imaginative exercise are adequate to the essential nature of the items for which a putative possibility is being entertained. The result is that metaphysical impossibilities may, in certain circumstances, be perfectly lucidly conceivable; and hence that the conceivability of pain's coming apart from any particular supposed physical identification of it is no indication of a genuine possibility for pain unless the phenomenal concept of pain is adequate to the nature of pain—which physicalism denies.

Keywords: [conceivability](#), [Hale](#), [Kripke](#), [physicalism](#), [possibility](#)

Subject: [Moral Philosophy](#)

Collection: [Oxford Scholarship Online](#)

It is a very special pleasure to contribute to this volume in celebration of the philosophy of my old friend and long-time collaborator, the late Bob Hale, and I am most grateful to Ivette Fred and Jessica Leech for creating an opportunity for us all to do so. I have vacillated somewhat in trying to select a topic. For over thirty years, from the time when Bob was preparing *Abstract Objects* for the Blackwell series, *Philosophical Theory*, which I then co-edited, he and I worked together in a more or less continuous, hugely enjoyable and instructive exploration of the problems, prospects, and possibilities facing the development of a philosophy of mathematics in the Fregean tradition, a project which has proved to have ramifications virtually all over philosophy. It would have been very natural to offer something on one of the multiplicity of themes of this agenda. In the end, though, I have selected a topic closer to the central issues of Bob's most recent book.¹ What follows is by way of a grateful comment on his characteristically searching critique² of an assessment I wrote some years ago³ of the famous challenge to physicalism about the mental that is outlined in the concluding pages of Saul Kripke's *Naming and*

Necessity. Like most of its commentators, Bob and I were agreed that Kripke's argument comes short. The interesting question is exactly *why* it comes short, a question that leads to the heart of one central problem about the epistemology of the metaphysical modalities.⁴

p. 267

1 The Counter-Conceivability Principle

Let us speedily review the background to Kripke's challenge. *Naming and Necessity* was a veritable cluster-bomb of original, interrelated, deep-reaching ideas. Among the marquee notions it introduced were included:

- The idea of rigid designation and the thesis that the proper names of natural language are, characteristically, rigid designators.
- The contention that the same is true of many common nouns including a large class of terms standing for natural kinds of things, states, and stuffs—'water', 'heat', 'tiger', 'elm', 'diamond', etc.
- The idea of natural kinds as identified by essential characteristics of which an ordinary competent understanding of words standing for them may be entirely innocent.
- The modern conception of the metaphysical modalities as grounded not in the ways we talk and think but in the natures of what we talk and think about.
- The consequent separation of the necessary from the *a priori* and resultant hospitality to a raft of *a posteriori* necessities.

It is a consequence of these proposals that any statement purporting to identify the essence of a natural kind will be, if true at all, necessarily so.⁵ The contention that 'water', e.g., rigidly designates a natural kind requires that the use of that English word (with its present meaning) in speaking of any scenario, actual or counterfactual, always denotes the same kind of stuff, just as the use of 'Richard Milhous Nixon' in speaking of any scenario, actual or counterfactual, always denotes the man who was actually president of the United States in 1970. Hence, as Putnam famously argued, (assuming that water is indeed a chemical kind) the proper description of a hypothetical situation where some colourless, tasteless liquid, naturally occurring in rainfall, lakes, and rivers, etc., but of a differing chemical constitution, displayed all the characteristic surface properties of water, would be not that in that situation some water would have a different chemical constitution to what water actually characteristically has, but that the lakes and rivers, etc., would not be filled with water. If 'water' rigidly designates a chemical kind and it is true that water is H₂O, then it is necessarily true.

p. 268

It immediately follows that any evidence *against* the necessity of such an identification is *eo ipso* evidence against its truth. But what should count as evidence against the necessity of such an identification in the first place? As I interpret Kripke, his discussion implicitly rests on a major assumption on which we should pause: that *all* purportedly metaphysically necessary statements, even those grounded in essences knowable only *a posteriori*, are hostage to what we can, to borrow Descartes' happy phrase, 'clearly and distinctly' conceive; that to accomplish what is genuinely a clear and distinct conception of a situation is to establish its possibility. This principle—what in Wright (2002) I called the *Counter-Conceivability Principle*⁶—invites us, of course, to provide an account of when a conception should rank as relevantly clear and distinct. But without taking that issue on directly, we can cash the principle's operational content as being this: that if one has what at least *appears to be* a lucid, detailed conception of how it might be that not *P*, then that should count as a good, albeit defeasible ground for its being possible that not *P*, and hence its not being necessary that *P*, whatever the subject matter of *P*.⁷

By the Counter-Conceivability Principle, then, all putative metaphysical necessities, even *a posteriori* ones, have to run the gauntlet of what we can, as we think, clearly and distinctly conceive. If we can indeed clearly and

distinctly conceive of a scenario in which the proposition that *P* would be false, then that will be to disclose, *a priori*, a defeater of the claim of *P* to record a metaphysical necessity, even if the sole possible form of justification for *P* would have to incorporate empirical information.

2 A Protest

The Counter-Conceivability Principle (henceforward CCP⁸) is apt to impress as a controversial assumption at best—and at worst, in the present context, as a crude mistake. Conceivings, even at their clearest and most distinct, are constrained only by the requirements of the *concepts* that they configure. The metaphysical-modal status of a proposition, by contrast, at least in the central cases where it is recognized *a posteriori*, is grounded in requirements imposed by the natures of non-conceptual *things*—objects, properties, and kinds, etc.—which its ingredient concepts are possibly inadequate, possibly incomplete, concepts of. Conceivings, at their best, may tell us what is and is not permitted by the concepts they work with; may tell us, in short, what is *conceptually* possible, impossible, or necessary. But Kripke's own work is exactly the *locus classicus* for the existence and importance of the distinction between the conceptual and the metaphysical modalities. The CCP, it may be charged, when harnessed to an attempt to engage the latter, implicitly surrenders that distinction. So any argument based on the CCP is dialectically unavailable to anyone who endorses the distinction.

That's the protest. There is surely something right about it. On the other hand, if claims of (unactualized) metaphysical possibility cannot be justified, even if defeasibly, by considerations of what is lucidly conceivable, how *can* they be justified? What is the epistemology of metaphysical possibility to be?

For the time being I am going to leave this issue as an elephant in the room. We'll come back to it.

3 The Anti-Physicalist Argument

p. 269

So, again: assume that water, the stuff, is indeed a natural kind, correctly identified at the level of Daltonian chemical theory as H₂O, and that 'water', the noun, is a rigid designator of that kind. Then

(i) Water is H₂O

is not merely true but necessarily true. Kripke anticipates⁹ the natural objection to this upshot: surely it is readily conceivable that water might have turned out *not* to be H₂O. We can easily imagine the science having worked out quite differently, even that water might have proved not to be any particular chemical kind at all but, as it were, a syndrome of surface properties with a multiplicity of variable underlying causes. But if the CCP is in force, then—since what might have turned out to be the case might thereby *have been* the case—that is a *prima facie* case that (i) is not necessary. And in that case we will have to deny on purely philosophical grounds that it is even true—which is absurd.

The objection is heavily invested in the CCP. One response for a defender of Kripke's modal views would be simply to drop the principle. Famously, though, that is not what Kripke did. Rather, he allowed that the objector may be lucidly conceiving *something* perfectly coherent but contended that what is so conceived is nothing to the purpose. What *is* conceivable is that a stuff that presents on the surface as water does should not be H₂O. But to conceive of a stuff that presents on the surface as water not being H₂O is not to conceive of *water's* not being H₂O.

Again, since both 'Hesperus' and 'Phosphorus' are, plausibly, rigid designators, the statement,

(ii) Hesperus is Phosphorus,

must, if true at all, be necessarily true. But it took quite painstaking astronomical observation and calculation to figure out that (ii) is true and it might seem readily conceivable that the enquiry might have turned out differently: that Hesperus, the brightest star-like body seen in the evening sky, might have turned out to be a distinct entity from Phosphorus, the bright 'star' characteristically visible low in the eastern sky just before sunrise. The Kripkean reply will then be, similarly: No, what is indeed conceivable is that a heavenly body that presented in the evening sky exactly as Hesperus does should have turned out not to be Phosphorus. But to conceive of a heavenly body that presented itself exactly as Hesperus actually does turning out not to be Phosphorus is not to conceive of *Hesperus*' turning out not to be Phosphorus.

p. 270 In sum: Kripke's response to protect the necessity of true identifications of the relevant kind from the depredations of apparently lucid counter-conceivings is to insist on a distinction between conceiving of *X* not being *F* and conceiving of an *epistemic counterpart* of *X* not being *F*. Something which presents as water in all (surface) respects covered by our pre-theoretic conception of water—the indicators ↴ that, prior to the scientific investigation, we would use to classify a sample as one of water—need be no more than an epistemic counterpart of water. Something that presents as Hesperus in all respects covered by the characteristics incorporated in our folk-identifications of Hesperus need be no more than an epistemic counterpart of Hesperus. If, in accordance with the CCP, an episode of imagination, however lucid, is to support a judgement of genuine possibility, it must be shown to engage with *X* itself, rather than a mere epistemic counterpart of *X*. Otherwise the claimed necessity of [*X* is *F*] cannot be brought thereby into question. The apparent counter-conceivability of [Water is H₂O], and of [Hesperus is Phosphorus], are *modal illusions*, spawned by this conflation.

The crux of Kripke's anti-physicalist argument is now, famously, that there is no such corresponding distinction, or available conflation, for pain. Pain has no epistemic counterpart that is not pain: whatever presents on the surface as pain (i.e. whatever hurts!) is pain. Using 'C-fibre-excitation' as parametric for any putative identification of pain in neurophysical terms, we thus have both that

(iii) [Pain is C-fibre excitation] is, if true, necessary,

—since both ingredient terms are, plausibly, rigid designators—and that

(iv) [Pain is C-fibre excitation] is counter-conceivable,

since it does seem perfectly possible fully lucidly to conceive of oneself, or others, being in pain in circumstances where their, or one's own, C-fibres prove, by appropriate scientific tests, to be inactive. And now there is no defending the purported identity by the charge that the imagination here deals merely in an epistemic counterpart of pain. Again, there *are* no mere epistemic counterparts of pain, no states that present in all respects as pains do but are not pains. Any epistemic counterpart of pain *is* pain; the imaginative episode *does* engage its target. So by the CCP, a case, so far undefeated, is made that pain is not necessarily C-fibre excitation. It follows within Kripke's assumptions that pain is not C-fibre excitation at all and hence, since the play with the latter was merely parametric, that pain is *no* type of physical state.

4 Illusions of Possibility: Another Putative Source

The most salient direction for a physicalist response to take is to pursue the thought that, granting that the apparent possibility of pain's not being C-fibre excitation cannot be a modal illusion in *that* kind of way, still maybe that is not the only possible template that an illusion of possibility may instance. Perhaps there are other ways whereby an apparently lucid conception of how *P* might be true can miss its intended mark—can fall short of an indication of *P*'s genuine possibility—other than by trading on a conflation between events, states, or stuffs, etc. and mere epistemic counterparts of them. Perhaps there are other ways whereby an apparently lucid conception of *P*'s obtaining can indeed actually be lucid and adequately detailed but be of something else, something that masquerades, as it were, as the obtaining of *P*.

In Wright (2002), I offered two putative examples to suggest one such alternative model. The first was the following fantasy:

p. 271

Suppose—Kripke would agree—that I am essentially a human being, and that it is an essential characteristic of human beings to have their actual biological origins. So it is an essential characteristic of mine to be the child of my actual parents. Still I can, it seems, lucidly conceive of my not having had those parents but others, or even—like Superman—of my originating in a different world, of a different race, and having been visited on Earth from afar and brought up as their own by the people whom I take to be my biological father and mother. I can, it seems, lucidly imagine my finding all this out tomorrow.¹⁰

Now, this example, I claimed, cannot be dismissed

on the ground that it fails to be sensitive to the distinction between myself and a mere epistemic counterpart, a mere 'fool's self', as it were, sharing the surface features by which I identify myself but differing in essence. It cannot be so dismissed because I don't, in the relevant fashion, identify myself by features, surface or otherwise, at all. The point is of a piece with Hume's observation that, in awareness of a psychological state as one's own, one is not presented as an object to oneself. When I conceive some simple counterfactual contingency—say, my being right at this moment in the Grand Canyon—I do not imagine someone's being there who presents themselves, on the surface, as being me. Rather I simply imagine *my* having relevant kinds of experience—imagine, that is to say, the relevant kinds of experience from my first personal point of view. No mode of presentation of the self need feature in the exercise before it can count as presenting a scenario in which *I* am in the Grand Canyon; a fortiori, no *superficial* mode of presentation, open to instantiation by someone other than myself.¹¹

The suggestion is that a sweeping range of lucid and detailed flights of fancy running counter to presumed metaphysical necessities are open to the first person—fancies of different biological origins, alternative modes of embodiment (Kafka's Beetle), even disembodied existence and conscious survival of the dissolution of one's body. And this is so precisely because first-personal thought is unconstrained by any mode of presentation of the self that stands to it as the surface indicators of water stand to the natural kind. Nothing in my concept of myself—the concept, that is to say, which I exercise just by the intelligent use of 'I'—obstructs the coherence of these fantasies. They can be as lucid and detailed as you like, yet are not open to diagnosis as missing their mark by virtue of working with mere epistemic counterparts. Still, the scenarios they construct appear to stand in flat conflict with widely accepted necessities of origin and constitution.

A different but broadly analogous challenge to the CCP, I contended, is provided by certain examples of apparent conceivability in mathematics. I wrote:

We can rest assured, I suppose, that Andrew Wiles really has proved Fermat's 'Last Theorem', which therefore holds good as a matter of conceptual necessity. But we can imagine a sceptic about the result who flatters himself that he can still conceive of finding counterexamples to the theorem, and of finding mistakes in Wiles' proof. Of course there will be limits on the detail of these or the sceptic would be thought-experimentally finding *real* counterexamples and mistakes. Still, we should not deny that he could be conceiving *something*, and doing so moreover—subject only to the preceding point—in as vivid and detailed a way as could be wished.

p. 272 About this case, however,

the last diagnosis we should propose...is that his conceivings are insensitive to the distinction between finding counterexamples to Fermat's theorem and finding counterexamples to an *epistemic counterpart* of it!—or to the distinction between finding a mistake in Wiles' reasoning and finding a mistake in an epistemic counterpart of that.¹²

To be sure, the issues pointed to by this example go beyond those raised by the question whether it is indeed conceivable that any particular identification of pain with a neurophysical kind should fail. For this time it is, plausibly, as suggested in the quoted passage, a *conceptual* impossibility which is presented as conceivable. The bearing of the example is directly on the tenability of the CCP even where it is conceptual possibility that is at stake. It accordingly raises very vividly the question, just what should we be asking of a relevantly 'lucid and detailed' conception. A natural objection to the case—tendered, indeed, by Hale—is that, while we may indeed grant that the sceptical mathematician is coherently conceiving something, the acknowledged unavoidable limitations to the detail of his conceivings rob the example of its intended force. The CCP, after all, speaks of clear and distinct—lucid and detailed—conceptions: a defender of the principle has no reason to deny that impossibilities may be *hazily* conceived. Well, no doubt. But must the conceivings of the sceptic in the example be *relevantly* hazy? There is no reason not to grant that whatever ideas our mathematician had, at a time before Wiles' discovery, both of what a proof of Fermat would have to be, and of what a counterexample would have to be, were definite enough to put him in position to recognize any instantiation of either, provided it could be surveyably written down—definite enough, indeed, to leave him with no space for discretion were he to be presented with such an instance of either. And presumably they are no less definite after the announcement of Wiles' result. But are not concepts that are definite enough to enforce recognition of any instance of them, with no foreseeable room left for discretion, thereby definite enough to provide the materials for conceivings that are relevantly lucid and detailed for the purposes of the CCP? If not, what is the standard to be?

These two cases, then, are *prima facie* relevantly lucid and detailed conceivings of impossibilities. The choices they present to a friend of the CCP are, accordingly, two: either—the *radical option*—to admit them and then to identify some suitable qualification of the principle that allows it still to be of some modal-epistemological use; or—the *conservative option*—to follow Kripke's example in trying to make a case that they do not, appearances notwithstanding, involve genuine conceivings of what they purport to be conceivings of.

My suggestion in Wright (2002) was of the latter, conservative kind. The examples indicate, I suggested, that there

has to be a category of conceivings that fall short of being counter-conceptions to a given proposition, not because their detail fails to be sensitive to the distinction between items that the proposition is about and 'fool's' equivalents of them

—epistemic counterparts—

p. 273

but because it is insensitive to another distinction: that between genuinely conceiving of a scenario in which *P* fails to obtain and conceiving, rather, of what it would be like if, *per impossibile*, *P* were (found

to be) false.¹³

On this proposal, there are examples not only of metaphysical but also of conceptual impossibility, *P*, such that we should not deny that one can perfectly lucidly and in detail conceive of scenarios that deserve the description: this is how things would/could be if *P* were to be false. What we deny is that such conceivings are tantamount to conceiving of *P*'s being false. The CCP is therefore spared. But at the same time we still provide the physicalist with the resources to write off the apparent counter-conceivability of [Pain is C-fibre activity] as a modal illusion. She may maintain that it is a modal illusion of this second kind: that is, she can allow that while it is apparently possible clearly and distinctly to conceive of pain in the absence of C-fibre excitation and, conversely, of C-fibre excitation in the absence of felt pain, it is not really possible. Rather, analogously to Kripke's explanation of the illusion of conceivability of water's not being H₂O, the physicalist can contend that while *something* is indeed lucidly and distinctly conceived in such flights of fancy, what is conceived is not the intended target. One conceives, it may be suggested, not of pain's not being C-fibre excitation, but—what is not the same thing—of what things would be like if, *per impossible*, it were something else.

5 Hale's Response

This response to Kripke's argument is not to Hale's liking. We should distinguish, he urges, between, on the one hand,

- (a) The question whether there are indeed other relevant species of illusion of possibility in addition to the species, highlighted by Kripke, involving play with mere epistemic counterparts of a targeted kind, or stuff, etc.,

and, on the other,

- (b) The question whether the putative *per impossible* species invoked in my treatment of the 'SuperWright' and 'Sceptical Mathematician' examples is either (i) needed to account for either of those cases, or indeed (ii) is anyway in good standing.

Hale thinks the answer to (a) is 'yes', and that the other relevant modes that he goes on to describe will sustain no less a negative appraisal of the anti-physicalist argument. But he has doubts about both (b)(i) and (b)(ii).

To focus first on (b)(ii). Hale reacts to the formulation in the passage from my (2002) most recently quoted above—the putative contrast between 'genuinely conceiving of a scenario in which *P* fails to obtain and conceiving, rather, of what it would be like if, *per impossible*, *P* were (found to be) false'. I rather regret that formulation insofar as it may encourage a reader to wonder—as did Hale—whether the impossibility marked by '*per impossible*' is to be an ingredient in the second kind of conceiving. The answer of course is that it is not: the second kind of conceiving is \hookrightarrow to be the source, after all, of a kind of *illusion* of possibility—so naturally there needs to be scope for someone to be taken in by it; the impossibility cannot be part of the explicit content of the conceiving itself.

Hale's principal doubt, however, about the *bona fides* of *per impossible* conceivings as characterized, focuses on the subjunctive conditional: 'what it would be like if...*P* were (found to be) false'. Against this he urges the obvious question: Where a proposition is impossible, would *anything* distinctively be the case if it were to hold?

Standard—Lewis-Stalnaker style—semantical treatments of counterfactuals say not, of course. In such treatments, a counterfactual conditional holds just if its consequent holds in all relevant worlds in which its antecedent does. Since, when a counterfactual has an impossible antecedent, there are no worlds—*a fortiori*, no relevant worlds—in which the antecedent does hold, there are no worlds in which the counterfactual fails to hold. So standard semantics declares such conditionals necessarily true, irrespective of their consequents. But

if all such conditionals are true, then there is nothing distinctive that the situation would be like if, say, the square root of 2 were rational or if I had been born of different parents. So there is nothing distinctive to be the content of the appropriate conceiving: you can conceive what it would be like if, *per impossibile*, P were to obtain, by conceiving anything at all.

Although I have no space to pursue the matter here, my own view is that, its theoretical naturalness notwithstanding, this consequence reflects badly on the standard semantical treatment of counterfactuals—that it indicates more about a degradation of content implicit in such treatments than about the putatively illusory character of any worthwhile distinction between, say, ‘If the square root of 2 were rational, it would be equivalent to m/n for some pair of integers, m and n ’ and ‘If the square root of 2 were rational, it would be equivalent to π ’. There is an interesting growing literature in sympathy with this reaction,¹⁴ though how to develop a semantics for counterfactuals to do better than the standard treatment in this regard remains controversial.¹⁵ I shall not pursue that matter further here. As things stand, it seems fair to grant Hale that the prospects for a successful response to the anti-physicalist argument by deployment of the *per impossibile* template for modal illusion have not yet been clearly made out.

May we do better by enlisting a further template of the conservative kind—one that stops short of the idea that examples like those considered involve genuine conceivings of the impossible? Hale suggests that we may and offers two further potential sources of illusion of possibility in particular. First on the Sceptical Mathematician:

p. 275

The sceptic starts with the hypothesis—in fact necessarily false, although he does not, of course, take it to be so—that there are counter-examples to FLT [Fermat’s Last Theorem]. He then imagines someone finding one—*without, however, attempting to imagine it in any detail*—and goes on to imagine the impact of this ‘discovery’ on the mathematical community: the initial reactions of shock and even disbelief which, however, quickly disappear (since \hookrightarrow simple and apparently flawless calculations show the putative counter-example to be the genuine article), the embarrassment of those (virtually everyone) who had been ‘taken in’ by Wiles’s proof, Wiles’s own chagrin and reluctant admission that there must be some flaw in his work, the ensuing hunt for a mistake—perhaps a hitherto unnoticed but false assumption—in that ‘proof’, and so on. There is, it seems, no doubt that this is a perfectly conceivable flight of fancy, and that it could be elaborated in as much vivid detail as you like, subject to the proviso that it remains silent on the mathematical detail of the supposed counter-example... What is right is that this is a perfectly intelligible piece of conceiving. What is wrong—or at least, in my view, very much open to question—is that the flaw in it, viewed as an attempt to counter-conceive FLT, should be diagnosed as Wright proposes, i.e. as a matter of insensitivity to a distinction between conceiving of FLT’s being false and conceiving of what it would be like if, *per impossibile*, FLT were (found to be) false. Nothing in what the sceptical mathematician actually *imagines*—as distinct from the background doubt as to the truth of FLT which *fuels* his imaginative flight—*demand*s its description in those terms, *rather than* in these quite different terms: conceiving what it would be like if, *as is perfectly possible*, FLT were...*thought* to be counter-exemplified.¹⁶

Now, I already argued that the level of necessarily missing mathematical detail in the Sceptical Mathematician’s imaginings should not immediately disqualify them from the status of adequately lucid and detailed conceivings in the spirit of the CCP. But I won’t pursue that further now. Hale’s proposal is that the example is better described—or anyway as well described—in terms of a *Doxastic Consequences* model. According to this proposal, we are apt to confuse conceiving of P ’s being (found to be) false with conceiving of what it would be like if, as is perfectly possible, P were *thought* to be false. And, Hale contends, nothing in the description of the case requires that we think of the Sceptical Mathematician as doing anything beyond what can be accommodated by the latter description.

We will consider that claim in a moment. But first here is Hale on SuperWright:

We should now come back to the question suggested just now, whether one should describe this kind of conceiving in terms of imagining (our) finding out or discovering that such-and-such, or rather in more guarded terms as imagining (our) having compelling evidence or good reason to believe that such-and-such. If Hume's Principle is true—i.e. if conceivability implies possibility—then we should insist upon the more guarded description. For if we really could conceive of finding out that *P*, then it would be possible that we should find out that *P*. And since 'find out' is factive—we can only find out that *P*, if *P*—it would follow (by the sound modal principle that $\Diamond A$ entails $\Diamond B$, when *A* entails *B*) that it is possible that *P*. It is true enough, or at least widely accepted, that in many—perhaps the vast majority of—cases, it is metaphysically, even if not epistemically, possible that the causes of events and states of affairs should have been other than they actually were. In such cases, we may agree that we can conceive of finding out that the causes were otherwise. But if some of the facts about causation are metaphysically necessary—as they must be, if the doctrine of essentiality of origin is true—then whatever it is that we can conceive in regard to causes being other than they actually were, it cannot be that we can conceive of *discovering* that they were other than they actually were. \downarrow The closest we can come to that is (something like) conceiving of our having compelling evidence, or good reason to believe, that the causes were otherwise.¹⁷

This reasoning seems to me compelling, as far as it goes. That is, if—as Hale is assuming—the CCP is good, that is, if genuine conceivability really does imply possibility, and *P* is necessarily false, then of course, given the factivity of 'finding out', we cannot genuinely conceive of finding out that *P*. In that case, we have no option but to redescribe in lesser terms whatever it is that is apparently conceived lucidly and in detail in the SuperWright example. Hale's suggestion is that the example exhibits what he calls *The Misleading Evidence model*—we think it possible that *P* because we can conceive of our having compelling evidence or good reason to believe that *P*, and we fail to distinguish this from conceiving of its being the case that *P*.

Now, I certainly don't want to deny that misjudgements of possibility *might* be spawned by either of the kinds of conflation that Hale draws attention to—by confusing how things would be if *P* with how things would be if we (all) believed *P* or if we had misleading evidence that *P*. But there is an immediate concern: it is not clear that either the Doxastic Consequences model or the Misleading Evidence model cuts any mustard when it comes to trying to explain away the appearance of the lucid and detailed conceivability of pain without C-fibre excitation or its converse. Someone who thinks—as most of us may be presumed to do—that one can vividly imagine what it would be like to be in pain absent any particular type of putative characteristic physical basis for it is most unlikely to be prepared to accept the watered-down suggestion: 'No: you are really just imagining what it would be like if you believed you were in pain while e.g. none of your C-fibres were firing', or 'No: you are really just imagining what it would be like if you had misleading evidence that you were in pain while e.g. none of your C-fibres were firing'. There is no clear sense to the idea of misleading evidence that one is oneself in pain. And there is no plausibility whatever in the suggestion that, in the imagined scenario, one is merely imagining consequences and surroundings of the belief that one is in pain. Not a bit of it: one imagines *pain*—the very phenomenal state itself, not a doxastic state and its surroundings and consequences. For one out to put a roadblock in the way of Kripke's anti-physicalist argument, Hale's two suggested additional models of misbegotten conceivings seem, so far, not to promise any leverage at all.

Hale, expectably, thinks otherwise. He writes:

In fact, Kripke's argument is in trouble anyway...so long as my *Doxastic Consequences* model is workable. For if that is so, then—whether or not it is agreed to be the best model for explaining the Sceptical Mathematician's modal illusion(s)—it could, clearly, be invoked by the identity-theorist to explain away the apparent conceivability of painless C-fibre [excitation] and C-fibreless pain. We can easily imagine how things might be, if pain were believed to be other than C-fibre [excitation]. We imagine, for example, media reports of neurophysiologists discovering that as well as C-fibres,

normal human beings have D-fibres. Since these are distributed about our bodies much as C-fibres are, it is readily understandable how it came to be thought that pain is C-fibre [excitation]. Analysis reveals, however, that whilst most ↪ pain-killing drugs reduce *both* C- and D-fibre activity, some newly developed ones reduce only the latter and are everywhere as effective as the usual drugs in reducing pain. Perhaps we go on to imagine reading newspaper articles reporting massive investment by the big pharmaceutical companies in research programmes aimed at developing newer and even better D-fibre ‘dullers’, and less well-funded investigations into the (now obscure) function of C-fibres, etc.¹⁸

But this train of thought would seem to be point-missing. The question is not whether we can in fact easily imagine what things might be like if pain were believed to be other than C-fibre excitation (even if it is essentially C-fibre excitation) but whether the characterization of what we are doing when we do that, outlined as by Hale, does justice to what we actually, as it seems to us lucidly and in detail, imagine when we imagine *pain without C-fibre excitation*. The fact is that it surely does not.

I envisage a protest. Surely, it may be suggested, this is to take Hale’s proposals the wrong way round: they will fare better for the physicalist’s purpose if we focus not on pain but on the second term in the putative identity, viz. C-fibre excitation (or whatever, parametrically, the physical basis of pain is taken to be). Let us grant the contrasts between having misleading evidence that one is in pain and really being in pain, or between merely believing that one is pain and really being in pain, struggle for any genuine phenomenological content. Still, the ideas of having misleading evidence of the absence of C-fibre excitation, as opposed to C-fibre excitation genuinely being absent, or of merely believing that there is no occurrent C-fibre excitation, as opposed to there really being none, do not, it is plausible, present the same difficulties. And if they do not, cannot Hale’s models be harnessed after all to help the physicalist account for an illusion of possibility in the relevant case? What, the physicalist may argue, is conceivable, is not pain without C-fibre excitation but pain in the presence of misleading evidence that there is no concurrent C-fibre excitation, or pain in circumstances where it is generally believed that there is no concurrent C-fibre excitation. Are not these ‘alibis’ as good as, though different in content to, the play with the idea of an epistemic counterpart marshalled in Kripke’s original defence of the necessary identity of water with H₂O?

I don’t think this improves matters. There is a dilemma. Is it or is it not to be granted that there is a difference in general between conceiving of an absence of C-fibre excitation in one’s occurrent neuro-physical make-up and conceiving of a situation where all the evidence misleadingly suggests such an absence or where it is at any rate accepted by all concerned that there is such an absence? If *these* differences are conceivable, then—however exactly the difference is accounted for—surely there can be no obstacle to one’s conceiving of the former state, rather than one of Hale’s surrogates. And then, since such a conception will presumably be entirely non-committal as far as one’s occurrent phenomenological state is concerned, there should be no obstacle to adjoining to it in imagination the felt experience of pain. Just that was exactly the nub of Kripke’s original argument. But if we say instead that the distinctions in question *cannot* feature in the content of an episode of conceiving—that while, of course, perfectly *intelligible*, they are, strictly, *inconceivable* distinctions—then, sure, pain without C-fibre excitation will ↪ now be inconceivable but so will be pain *with* C-fibre excitation and pain, or anything else, with, or without, any of the range of circumstances which have now implicitly been pushed beyond conceivability. I won’t here speculate by what principle that range might be determined, but one would suppose that once the distinction between C-fibre excitation and misleading evidence of C-fibre excitation is held to be ineligible to feature within the content of any relevant conceiving, the extent of the parallel casualties thereby entrained will be very wide indeed—wide enough to put a huge sweep of metaphysical possibility claims beyond the reach of corroboration by lucid and detailed conceiving. The cost of protecting the necessity of the putative identity of pain with C-fibre excitation in this way is thus utterly to emasculate the CCP.¹⁹

There is one more move to consider. We may continue to allow that the contrast between conceiving of an absence of occurrent C-fibre excitation and conceiving of a situation in which we merely have misleading evidence of such an absence, e.g., can still be active in the description of the content of a subject's conceivings, but that the distinction is *phenomenologically* underdetermined: that which characterization is to be preferred will turn, rather, on the modal status of the proposition whose possibility, when the episode of conceiving is described in the former way, it would, via the CCP, support. So if pain is necessarily C-fibre excitation, then the characterization of the relevant episode of conceiving as being of pain without C-fibre excitation will be illicit. But if pain is not necessarily C-fibre excitation, then that characterization can stand—even though subjectively the episode in question may impress as exactly the same.

I won't need to dwell on the obvious methodological difficulty with this. The CCP can have no part to play in an interesting epistemology of the modal if whether anyone counts as entertaining an adequately lucid and detailed conception of circumstances in which *P* would hold is a question that cannot in general be answered positively without an independent assessment of *P*'s modal status; if the distinction, that is, between a genuine and a merely apparent counter-conception cannot in general be drawn purely by reference to the actual detail of the conceivings concerned. In general, if the CCP is to hold as a *basis for judgement*, then warrant for a conceivability claim has to be available independently of collateral assessment of the actual modality of the proposition concerned. The distinction between genuinely conceiving [not-*P*] and merely conceiving [*P**], somehow confused with [not-*P*], has to be available, and operational, prior to any verdict on the modal status of *P*.

6 Conceiving the Impossible

p. 279

It is high time to pay attention to our neglected elephant. Let us note, to begin with, that the force of the SuperWright and Sceptical Mathematician examples does not depend on their being presented under the *per impossibile* template—i.e. as cases where a lucid and relevantly detailed conception succeeds in being a conception not of the truth of *P* but—what need not amount to that—only instead of how things \hookrightarrow would be in certain respects if *P*, *per impossibile*, were true. The issues about the status of counterfactual conditionals with impossible antecedents raised by so presenting the examples may accordingly be sidestepped. The challenge to Kripke's argument that they pose will hardly be diminished if they may be taken to illustrate how it can be *prima facie* possible to conceive in a relevantly lucid and detailed manner of the *actual* obtaining of what are, in metaphysical fact, metaphysical impossibilities. Both the SuperWright and Sceptical Mathematician examples can be easily addressed to this purpose: *re* the former, it doesn't detract from the credentials of the relevant play in the imagination if I am deluded enough to suppose that what I imagine about my origins and early history is *actually* true; *re* the latter, it was already a feature of the example that the Sceptical Mathematician does exactly that. Provided such *prima facie* conceivings resist deflation by the Kripke style of explanation in terms of epistemic counterparts, they will present an unanswered challenge both to the CCP and to its annexure to the anti-physicalist argument.

To receive the examples in this way is, of course, to take the *radical* option distinguished earlier:²⁰ it is to regard them as so far undefeated counterexamples to the CCP, and hence to be prepared to allow that conceivability itself can be a source of modal illusion in a fashion contrasting with any model that postulates that in illusory cases our conceivings miss their mark—that we conceive of something *other than* the truth of the targeted proposition. I am suggesting, therefore, that we should abandon the CCP-conservative approach to modal illusion shared by the epistemic counterpart, *per impossibile*, misleading evidence and doxastic consequences models and take seriously the salient doubts, voiced in Section 2, about the standing of the CCP itself. Conservative approaches attempt to maintain that lucid and detailed conceivability—*real* lucid and detailed conceivability—suffices for metaphysical possibility. So where *P* is impossible yet we have what appears to be a lucid and detailed conception of *P*'s obtaining, that conception must either be defective in some way or be a

perfectly good conception of a surrogate scenario which we somehow mistake for the obtaining of *P*. Yet if conceivability is to supply a sure criterion for metaphysical possibility, this requires that in any case where an apparently lucid and detailed conception seems to be available of the obtaining of what is in fact an impossibility, we are guaranteed to be able either to find relevant fault with its credentials to be lucid and relevantly detailed or to make a good case that it deals with a surrogate. I suggest that, even over the short space of the preceding discussion, a serious question has emerged whether we can in general redeem any such guarantee. No: when it is metaphysical possibility that is at issue, why not just recognize that some metaphysical impossibilities may be perfectly lucidly conceivable—precisely because the impossibilities concerned are not grounded in the first place (purely) in our *concepts* of the events, states, or stuffs, etc. concerned?

p. 280

We noted the outstanding objection to this move: that to abandon the CCP is to risk leaving ourselves with no effective epistemology of metaphysical possibility—no means for properly controlling judgements of metaphysical possibility save when they are grounded in actuality. This, though, should impress as an overreaction. ↪ For one thing, provided we retain some form of effective control on ascriptions of metaphysical *necessity* (and thereby on impossibility, the necessity of a contradictory), it is not clear that we anyway also need an effective epistemology of possibility. Rather, the metaphysical possibility of *P* can be treated as a default status, defeasible by grounds for the necessity of some *Q* that is inconsistent with *P*.²¹ But for another it is in any case still open to us to treat lucid and detailed conceivings as providing positive though defeasible grounds for possibility claims. What is required is only that we acknowledge that they can misfire in either of two general ways:

- (i) First, as grounds for *conceptual possibility* claims, they can misfire if they suggest the possibility of something that is indeed conceptually impossible but for reasons sufficiently subtle to lie well beyond anything whose recognition is required by an ordinary competent, reflective understanding of the concepts concerned. (The Sceptical Mathematician)
- (ii) As grounds for *metaphysical possibility* claims, lucid and detailed conceivings can misfire if they work with concepts that misrepresent,²² or are silent on, aspects of the metaphysical nature of the objects they concern. Misrepresentation may result in the apparent exclusion of genuine possibilities; silence in the recognition of spurious possibilities.

The latter case gets us something close in spirit to Kripke's diagnosis of the water = H₂O case, even though our outlook now on the CCP is radical rather than conservative; for we are enabled to diagnose the case not indeed as involving imaginative play with an epistemic counterpart of water but as working with an incomplete concept of water. Simply, the ordinary concept of water is silent on the essential nature of water. So it allows for the detailed and lucid imaginability and thereby the conceptual possibility of scenarios in which water is other than H₂O. The conclusion that so much is metaphysically possible would indeed be licensed if it could be assumed that the concepts exercised in the imaginative scenario were fully indicative of the essential nature of the stuff. However, they aren't. There is no need to say that what is conceived is merely that an epistemic counterpart of water is not H₂O. What is clearly and lucidly conceived is, rather, the conceptual possibility of a metaphysical impossibility.

Crucially, however, unlike the diagnosis in terms of imaginative play with mere epistemic counterparts, the same form of debunking explanation may be applied smoothly to SuperWright. One who subscribes to the (widely accepted) essential animal nature of human selves, and the necessity of our actual biological origins, should insist that the concept of the self expressed by 'I' that is active in the flights of fancy of the kind concerned is simply silent on these matters. It therefore does indeed generously allow of conceptual possibilities that transcend the metaphysical possibilities. The fancy of being born of other parents, etc. is such

a possibility, lucidly imaginable in detail without conceptual incoherence, but metaphysically impossible nonetheless.

7 Coda: The Anti-Physicalist Argument

It should be clear now what a physicalist must say in response to Kripke's argument. She must counter that any purely phenomenal concept of pain is similarly inadequate to the essential nature of pain. Pain, for a plausible physicalism, must have a *dual essence*, with ineliminable phenomenal and physical components.²³ The apparent conceptual possibilities opened up by concentration on just one of these are or can be genuine conceptual possibilities, imaginable lucidly and in detail, and are properly described as realizations in imagination of conceptual possibilities for pain. But the conceptual possibilities for pain are a broader class than the metaphysical possibilities, and such flights of fancy bear on the latter only given the—for the physicalist—erroneous assumption that the phenomenal concept of pain is adequate to its essence.

'But how can pain have a *dual essence* of this kind?' Well, that was a conundrum implicit in physicalism from the start. In effect, it is just the form in which the so-called Hard Problem of Consciousness presents itself to physicalism. It is, in one shape or another, a problem for everyone.²⁴

References

Berto, F. (2014) 'On Conceiving the Inconsistent', *Proceedings of the Aristotelian Society* 114, 103–21.

[Google Scholar](#) [WorldCat](#)

Berto, F. and Schoonen, T. (2018) 'Conceivability and Possibility: Some Dilemmas for Humeans', *Synthese* 195(6), 2697–715.

[WorldCat](#)

Boyd, R. (1992) 'Materialism without Reductionism: What Physicalism Does Not entail' in N. Block, ed., *Readings in the Philosophy of Psychology*, Cambridge, MA: Harvard University Press, Vol. I, pp. 67–106.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Brogaard, B. and Salerno, J. (2013) 'Remarks on Counterpossibles', *Synthese* 190(4), 639–60.

[Google Scholar](#) [WorldCat](#)

Divers, J. (2002) *Possible Worlds*, London: Routledge.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Hale, B. (2013) *Necessary Beings: An Essay on Ontology, Modality, and the Relations Between Them*, Oxford: Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Hale, B. (forthcoming) 'CCCP' in A. Miller, ed., *Logic, Language and Mathematics: Themes from the Philosophy of Crispin Wright*, Vol. 2, Oxford: Oxford University Press.

Hume, D. (1888) *A Treatise of Human Nature*, ed. L.A. Selby-Bigge, Oxford: Clarendon Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Kment, B. (2006a) 'Counterfactuals and Explanation', *Mind* 115, 261–310.

[Google Scholar](#) [WorldCat](#)

Kment, B. (2006b) 'Counterfactuals and the Analysis of Necessity', *Philosophical Perspectives* 20, 237–302.

[Google Scholar](#) [WorldCat](#)

Kripke, S. (1980) *Naming and Necessity*, Cambridge, MA: Harvard University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Kung, P. (2010) 'Imagining as a Guide to Possibility', *Philosophy and Phenomenological Research* 81, 620–63.

[Google Scholar](#) [WorldCat](#)

Kung, P. (2014) 'You Really Do Imagine It: Against Error Theories of Imagination', *Nous* 50, 90–120.

[Google Scholar](#) [WorldCat](#)

p. 282 Nolan, D. P. (1997) 'Impossible Worlds: A Modest Approach', *Notre Dame Journal for Formal Logic* 38, 325–527.

[Google Scholar](#) [WorldCat](#)

Nolan, D. P. (2013) 'Impossible Worlds', *Philosophy Compass* 8: 360–72.

[Google Scholar](#) [WorldCat](#)

Wright, C. (2002) 'The Conceivability of Naturalism' in T. Gendler and J. O'Leary-Hawthorne, eds, *Imagination, Conceivability and Possibility*, Oxford: Clarendon Press, pp. 401–39.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Notes

- 1 Hale (2013).
- 2 Hale (forthcoming).
- 3 Wright (2002).
- 4 I should advise the reader that while I shall eventually offer some further reflections about how physicalism may defend against Kripke's argument, it will be (his tacit assumptions about) the epistemology of metaphysical modality that will be my main concern here. I will not, for the most part, be further exploring other issues to which Kripke's argument gives rise—for instance whether his argument, even if successful as given, can engage token-token forms of physicalism; whether natural kind terms are indeed, characteristically, rigid designators: whether, if so, their reference has to be construed at a deep, rather than a superficial level; what indeed rigidity comes to once singular reference and singular terms are left behind. These are all questions which a full appraisal of the argument would demand we look at, and Wright (2002) does have things to say about them. Here, though, our primary focus will be simply as my title suggests. And while I shall be drawing on some of the details of Bob's and my earlier exchange, I have done my best to presuppose no acquaintance with that.
- 5 At least, it is a consequence provided the characterization of the essence appealed to (namely, in the stock example, 'H₂O') is itself a rigid designator.
- 6 'Counter-Conceivability' as a reflection of the envisaged use of the principle against claims of metaphysical necessity.
- 7 The Counter-Conceivability Principle is something which might really have *deserved* the title of 'Hume's Principle', now of course purloined by the neo-Fregean programme for the foundations of arithmetic. Recall *Treatise* Bk. I, pt. II, section II: '*whatever the mind clearly conceives, includes the idea of possible existence, or in other words, that nothing we imagine is absolutely impossible*'.
- 8 I am going to drop Bob Hale's third 'C'!
- 9 The discussion of this point in Lecture III of *Naming and Necessity* actually proceeds in terms of other examples—Hesperus and Phosphorus, heat and molecular motion. So there is an element of interpretation in the assumption that Kripke would have endorsed the account to follow.
- 10 Wright (2002, 435).
- 11 Wright (2002, 435–6).
- 12 Wright (2002, 436–7).
- 13 Wright (2002, 437).
- 14 See for instance Kment (2006a, 2006b), Kung (2010, 2014), Brogaard and Salerno (2013), and Berto and Schoonen (2018).
- 15 For proposals, see Berto (2014), and Nolan (1997, 2013).
- 16 Hale (forthcoming).
- 17 Hale (forthcoming).
- 18 Hale (forthcoming).
- 19 Compare the discussion of Richard Boyd's (1992) response to Kripke's argument in Wright (2002) at pp. 413–17.
- 20 See p. 272.
- 21 The idea that we need no dedicated epistemology of metaphysical possibility is a feature of the discussions of both Divers (2002) and Hale (2013).
- 22 I am assuming that our concept of something may misrepresent it in certain details yet still succeed in referring to it, i.e. succeed in being of the item concerned.
- 23 This can be a stance available for type-type physicalism, though the physicalist may do well to leave room for a variety of essential physical types of pain, some perhaps species-wide (but maybe no wider), others tied to varieties in our phenomenal feels.
- 24 I presented the central argument of this chapter at the Diaphora workshop on the *A Priori* held at Stirling on 6–8 September 2017 with the support of the European Research Council. Thanks to the participants at the workshop for helpful discussion. I am also grateful to Paul Boghossian for useful critical comments on an earlier draft. Above all, especial thanks to Bob Hale.