

The Unsoundness of Arguments From Conceivability

Andrew Bailey
Department of Philosophy
The University of Guelph
Guelph, Ontario N1G 2W1
Canada
(519) 824-4120 x3227
abailey@uoguelph.ca

14 June 2007

ABSTRACT

It is widely suspected that arguments from conceivability, at least in some of their more notorious instances, are unsound. However, the *reasons* for the failure of conceivability arguments are less well agreed upon, and it remains unclear how to distinguish between sound and unsound instances of the form. In this paper I provide an analysis of the form of arguments from conceivability, and use this analysis to diagnose a systematic weakness in the argument form which reveals all its instances to be, roughly, either uninformative or unsound. I illustrate this conclusion through a consideration of David Chalmers' modal argument against physicalism.

The Unsoundness of Arguments From Conceivability

It is widely suspected that arguments from conceivability, at least in some of their more notorious instances, are unsound. However, the *reasons* for the failure of conceivability arguments are less well agreed upon, and it remains unclear how to distinguish between sound and unsound instances of the form. In this paper I provide an analysis of the form of arguments from conceivability, and use this analysis to diagnose a systematic weakness in the argument form which reveals all its instances to be, roughly, either uninformative or unsound. I illustrate this conclusion through a consideration of David Chalmers' modal argument against physicalism.

Here is a template for a *prima facie* useful, and widely encountered, type of a priori argument:

- i) There is a distinction to be made between ideal, or real, and merely apparent conceivability.
- ii) Ideal conceivability is a reliable indicator of logical possibility: that is, in general, if something is ideally conceivable then it is logically possible, and if it is not ideally conceivable then it is not logically possible.
- iii) X is (not) ideally conceivable.
- iv) Therefore X is logically (im)possible.
- v) Some claim about the actual world follows from the modal claim made in iv).
For example, if the non-existence of some entity is logically impossible, this establishes that this entity (say, God) must actually exist. Or if two things are logically possibly non-identical, then they (say, mind and body) are actually non-identical.

I will call arguments of this form 'arguments from conceivability.' They function by moving from an epistemic claim (about what can be conceived) to a modal claim (about what is logically necessary or possible) to a metaphysical claim about the nature of things in the world. Examples of this kind of argument are commonly found in debates about whether the mental is reducible to the physical (including most famously Descartes'

argument against mind-brain identity), or about the status of putatively necessarily existent entities such as God (perhaps most notoriously the ontological argument), or about certain sorts of reductive conceptual analysis (such as the attempt to analyse causal laws in terms of regularities in nature, or knowledge in terms of justified true belief, or goodness in terms of utility-maximization, and so on).¹

In any argument from conceivability there are two premises that, on the face of it, need special defence: in my characterization of this argument form, these are premises ii) and iii). Premise ii) is the claim that ideal conceivability is a reliable indicator of logical possibility. The most famous source of this principle is Hume: it is, he says, “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*” (*Treatise* 1.ii.2). However, prima facie this is a rather dubious claim. Conceivability is, after all, a subjective, psychological and epistemic property, while genuine (rather than merely apparent) logical possibility is usually taken to be a mind-independent, modal property, and it is not clear how this gap is to be bridged.

This, however, will not be my complaint in this paper: although I’m inclined to agree that it is mysterious how to make the leap from the epistemic to the modal, I think that it must somehow *be* possible or we couldn’t know anything about possibility and necessity at all. (After all, what other epistemic avenues have more prima facie plausibility? Empirical evidence from the actual world? Raw modal intuitions? I shall discuss this question in a little more detail below.) Ideal conceivability is apparently our main guide to logical possibility, and if it doesn’t work then what follows is a kind of generalized modal scepticism. There are various careful—if more or less tentative—defences of this jump across the epistemic-modal hurdle in the recent literature (for example, Yablo 1993, Levine 1998, and Chalmers 2002) and I defer to these.

The more fruitful question to ask, and the one I want to dwell on here, is something like the following: *under what circumstances* is it appropriate to argue from an epistemic or psychological claim to a modal one? Even if we accept that sometimes moves of this sort

can be made, clearly they are not always sanctioned: the inability of our ancestors in the fourteenth century to imagine, say, genetic engineering does not show that genetic engineering is impossible, any more than their belief that they could conceive of lead being transmuted into gold by the application of the appropriate chemical process (consistently with all the actual laws of physics) showed that this is in fact logically possible. Questions of this sort amount to demands for the defence of premise iii): the claim that X is possible depends upon the claim that it is conceivable in *the right way*—what I have called “ideally conceivable”—and premise iii) asserts that it is.

In Part 1 I will lay out the underlying form of arguments from conceivability and present my critique; in Part 2 I will consider objections and replies to this critique; and in Part 3 I will apply this critique to a recent, influential argument from conceivability, David Chalmers’ argument that phenomenal consciousness is non-physical.

PART 1

What would it be to show that X is ideally conceivable? I take it that an uncontroversial, minimal characterization of ideal conceivability would be something like the following:

X is ideally conceivable (if and?) only if someone who has adequate cognitive capacities and who is in possession of all the relevant factual and linguistic background information, and who has no additional distorting beliefs, could conceive of X.

That is, roughly, X is ideally conceivable only if it is conceivable by someone in relevantly ideal epistemic circumstances. Minimally, the conception of someone who was subject to relevant epistemic limitations could not be said to warrant a move to the modal pivot of an argument from conceivability *if* that conception would be different in a possible world where those limitations were removed.

Clearly there is far more to be said about ideal conceivability than this. Just to flesh out this definition, we would have to say something about ‘adequate cognitive capacities,’

‘relevant background information,’ and ‘additional distorting beliefs.’ I won’t attempt anything like a full characterization here, but I will give a brief illustration of what might be meant by these conditions.² A dog, presumably, does not have adequate cognitive capacities to test the limits of logical possibility—that no dog can imagine a prime number between 1000 and 2000, for example, says nothing against its possibility.³ In order to judge whether it is physically impossible to travel faster than the speed of light we need to know what the laws of physics are; this is a straightforward example of what I mean by relevant background knowledge. And an example of an additional distorting belief might be the case of a neurotic gambler who firmly but falsely believes that God will never let him win the jackpot, despite that fact that he sees other people win the jackpot all the time.⁴

No matter how these details are fleshed out, my main point for present purposes is that the following is a necessary condition on any plausible account of ideal conceivability: one’s conceiving of X can count as ideal *only if* the *removal* of any existing epistemic distortions (of the general sort described above) would *not* result in X’s ceasing to be conceivable. For example, if we *could* come to see X as inconceivable through acquiring a new piece of knowledge, then X cannot *now* be for us *ideally* conceivable.

This is not by any means an overly ambitious criterion for ideal conceivability. In particular, it does not require that some X can only be for us ideally conceivable if we are *already* in relevantly epistemically ideal conditions—a tall order indeed. It merely requires that if we *were* in relevantly epistemically ideal conditions we would continue to see X as conceivable. Indeed, this condition is simply a requirement of the fact that ideal conceivability is supposed to be a reliable a priori indicator of logical possibility: this requires minimally that its judgements should be *stable* in the face of the acquisition of new knowledge, or the shedding of false beliefs.

But now we are in a position to note the following: to assert that something is ideally conceivable is to make what is partly a *modal* claim—in order for something to be ideally conceivable it must be conceivable under the right *counterfactual* conditions. Thus, modal

judgements which are supported by arguments from conceivability inevitably tacitly rest on *another* modal claim: modal claim p is true if the situation described in p is ideally conceivable, but this situation is ideally conceivable *only* if it would still be conceivable under relevantly ideal epistemic conditions. If this latter modal claim, q , is false, then p remains unsupported.

So, a full justification of p on the basis of an argument from conceivability will require the establishment of q , another modal claim.⁵ If it is right that modal claims generally—or at least, contested modal claims—are to be established on the basis of arguments from conceivability, then we will have to provide *another* argument from conceivability for q .⁶ That is, we can ask whether we can imagine a possible world in which we have adequate cognitive capacities and all the relevant background information and no distorting beliefs, and in which X is conceivable. But *this* argument from conceivability will also depend upon an appeal to ideal conceivability—the relevantly epistemically ideal possible world in which X is conceivable must *itself* be ideally conceivable—and this in turn will require another modal claim which will require independent support, and so on.

In general, consider an argument form R and suppose that it were the only way of establishing propositions of the type s . Now suppose that one of the premises required for any argument of form R had to be of type s . It would follow that to establish the soundness of argument r_1 one first had to demonstrate the truth of proposition s_1 and that this would require an argument r_2 . But r_2 , since it is of the form R , will require for one of its premises the claim s_2 ; this premise in turn can only be established by an argument r_3 , of type R , and so it goes on. Hence, unless we reach the point where some premise s_n is so self-evidently true that it requires no further justification, the conclusion s (for any argument of type R) can never be finally established.

The systematic problem with arguments from conceivability, then, is the following: unless we are already in relevantly epistemically ideal conditions, the justification of the modal sub-conclusion of an argument from conceivability can *in principle* never be completed. In particular, the premise that asserts the ideal conceivability of X can never be justified and so

arguments from conceivability can never be established to be sound. It is not merely that the conceivability premise is, for any given instance of the argument form, incompletely justified: it's that for every instance of the argument form this premise can *never* be justified *at all*. To put it another way, for some contested modal conclusion p to be justified, an argument from conceivability A must be established as sound; but argument A can only be shown to be sound by another argument from conceivability, B. Since B is also an argument from conceivability, it in turn can only be established by argument from conceivability C, and so on in an infinite regress. This regress is vicious since p is not justified unless all the premises of A are, but all the premises of A are not justified unless all the premises of B are, and so on: since the regress is infinite, the justification of p is never—even partially—achieved.

PART 2

The point here is not merely the familiar one that a justificatory sequence must end either in some set of epistemic foundations or be resolved into a mutually justifying web of coherence. The problem for arguments from conceivability is much more pointed than this: the trouble lies in the fact that a given argument from conceivability can only be established on the basis of a further, more fundamental argument *of the same form*; this means that a regress can in principle never be completed. Furthermore the kinds of regress in question, clearly, will never circle back to the original claim, and this prevents a mutually supporting but individually unjustified set of conceivability arguments from being assembled.

Along similar lines, one might wonder whether the attack on arguments from conceivability I am presenting here depends essentially on an internalist model of justification: that is, am I tacitly assuming that to be justified in a modal belief one must have *reasons* for that belief? In a way I am, but I think not illicitly (and certainly without making a blanket commitment to internalism in epistemology). First, externalism with respect to *modal* beliefs is not, on the face of it, at all plausible, since there are no relevant analogues in the case of possible worlds to the causal and historical connections we have to the *actual* world

which typically underwrite reliabilist epistemologies.⁷ Second, the topic at issue in this discussion is not so much the question of when, if ever, we come to *know* some modal fact but the question of how, when called upon to do so, we can *defend* or *justify* a claim to know some modal fact. To do *this*, whether or not one adheres to an internalist account of justification in general, the challenged party must present an argument and, furthermore, must be able when called upon to show that this argument is sound ... and it is precisely this that the proponent of an argument from conceivability will never be able to do.⁸

One can sum up this discussion in the following way: arguments from conceivability as a class⁹ are either ‘trivial’—in the sense that their premises require no further justification—or unpersuasive, in that they cannot be shown to be sound. The modal sub-conclusions of arguments from conceivability fall most clearly into the former category when it can be established, or at least plausibly assumed, that we are *already* in relevantly ideal epistemic circumstances.¹⁰ However, in cases where we may well not be in relevantly ideal epistemic circumstances—cases which involve conceiving of a maximally perfect being, for example, or of a completed final physics—appeals to conceivability are rationally unconvincing in principle. I will bring this paper to a close by giving a short illustration of this dilemma in action.

PART 3

Perhaps the most famous conceivability argument in the recent philosophical literature is David Chalmers’ argument against physicalism, found in his book *The Conscious Mind* (1996) and in various other articles. Chalmers summarises the main thread of his book as follows (1996, 123):

- i) In our world, there are conscious experiences.
- ii) There is a logically possible world physically identical to ours, in which the positive facts¹¹ about consciousness in our world do not hold.
- iii) Therefore, facts about the consciousness are further facts about our world, over and

above the physical facts.

iv) So materialism is false.

A proper subset of the possible worlds appealed to in premise two—those which are physically identical to ours and yet different with respect to the distribution of consciousness—are sometimes called ‘zombie worlds.’ These, as the name suggests, are worlds physically identical with ours but in which consciousness is entirely *absent*. For the sake of concreteness, I shall focus on the case of zombie worlds. Chalmers’ argument, then, rests on two claims:

1. “If a physically identical zombie world is logically possible, it follows that the presence of consciousness is an *extra* fact about our world, not guaranteed by the physical facts alone” (1996, 123).
2. Physically identical zombie worlds *are* logically possible.

What makes Chalmers’ argument an argument from conceivability is that his defence of *both* of these claims rests ultimately on considerations about what is and is not conceivable. In order to establish Claim 1—that the logical or conceptual possibility of zombie worlds is sufficient to falsify materialism—Chalmers must answer philosophers who claim that materialism is content to rule out *metaphysically* possible zombie worlds, and is *consistent* with the ‘logical possibility’ of zombie worlds. In other words, Chalmers must show that the relevant modal judgements are *a priori* rather than *a posteriori*.¹² And in order to establish Claim 2, Chalmers must show that zombie worlds are in fact logically possible.

First, a word about Chalmers’ defence of Claim 1. The distinction between logical and metaphysical possibility is generally taken to rest on Kripke’s introduction of *a posteriori* necessities in “Naming and Necessity.” That is, what is logically (often called, in this context, epistemically) possible is what is possible *a priori*, and what is metaphysically possible is the (presumably smaller) class of things which are possible *a posteriori*. For example, in 1750 it was coherently imaginable that water not be H₂O—it was not, for example, an analytically false or formally self-contradictory claim. In this sense, then, the claim that water is not H₂O

is *a priori* logically possible. However, Kripke famously argued, since water *is* actually H₂O (and since all identities in which both terms are rigid designators are necessary identities) then water is *necessarily* H₂O and so the sentence “water is not H₂O” is an example of an *a posteriori* impossibility: it turns out, according to Kripke, that there are no *metaphysically* possible worlds where this identity fails to hold. And just as the *a priori* conceivability of water not being H₂O does not show that water is not identical with H₂O, it seems that the mere fact we can imagine zombie worlds need not falsify materialism. In other words, we might say, although materialism is false if there are possible zombie worlds, *a priori* considerations to do with conceivability are by themselves insufficient to establish the objective possibility of zombie worlds.¹³

Chalmers, however, argues that Kripkean examples such as these can all be accounted for in a framework—which he calls “two-dimensionalism,” or sometimes “modal rationalism”—in which the distinction between logical and metaphysical necessity falls at the level of statements and not of worlds. That is, on Chalmers’ account, judgements of *a posteriori* possibility do not place restrictions on sets of possible *worlds* but instead demarcate between evaluations of *statements* over the set of all possible worlds—that is to say, over the set of logically possible worlds. The significance of this, according to Chalmers, is that Kripkean considerations about *a posteriori* necessity fail to establish that materialism is compatible with zombie worlds: according to Chalmers, there is no possible world in which zombies are metaphysically impossible and logically possible, and so if we can produce a logically possible zombie world it must also be a metaphysically possible world. In other words, our epistemic access to the space of possible worlds is independent of *a posteriori* facts.

This is not the place to go into details about Chalmers’ two-dimensionalism. Very briefly, Chalmers distinguishes between two kinds of intension. Primary intensions are functions from possible worlds to referents which reflect the way actual-world reference is fixed; for example, the primary intension of “water” is a function to all the watery-stuff in actual and

possible worlds. Secondary intensions are functions from possible worlds to referents *once the primary intension has been rigidified*; thus, the secondary intension of “water,” across possible worlds, is a function to H₂O. Reference to water in other possible worlds on its *primary* intension requires no *a posteriori* knowledge about how the world turns out: there is thus, Chalmers says, a logically possible world in which water is not H₂O, and this is knowable *a priori*. By contrast, as Chalmers puts it, “*a posteriori* necessities are statements with a contingent primary intension and a necessary secondary intension. It is relatively easy to see that necessities of this sort cannot save the materialist” (1999, 477).

For the purposes of this paper, I simply want to emphasise the following point which is merely about the *structure* of Chalmers’ argument, rather than its content: in collapsing together the space of metaphysically possible worlds and the space of logically possible worlds, Chalmers places the whole weight of his argument on the *a priori* claim that zombie worlds are logically possible. And Chalmers argues for the logical possibility of zombie worlds by arguing that they are ideally conceivable, and asserting that something is logically possible if and only if it is ideally conceivable.

Chalmers’ argument for property dualism, then, has the following structure:

- i) The space of logically possible and metaphysically possible worlds is the same: that is, a world (rather than a statement) is metaphysically possible just in case it is logically possible.
- ii) Therefore, if a zombie world is logically possible it is metaphysically possible, and so if a zombie world is logically possible then materialism must be false, on any plausible view of materialism.
- iii) What it is to be a logically possible world is to be a “conceptually possible” world, where “conceptual possibility” is defined as: “conceivable on ideal rational reflection,” or “ideally conceivable” (1996, 35).
- iv) Therefore, if zombie worlds are ideally conceivable, they are logically possible and materialism is false.

- v) Zombie worlds are ideally conceivable.
- vi) Therefore, materialism is false.

Chalmers then proceeds to defend premise v) in three different ways, presenting five arguments in its support. First, he argues directly that certain situations where qualia are redistributed over the physical are *conceivable*: here he cites the logical possibility of zombies (1996, 94–99), and of inverted spectra (1996, 99–101). Second, he argues that *knowledge* of physical facts cannot entail the right kind of knowledge of phenomenal facts, and brings up the epistemic immediacy of qualia (1996, 101–103) and Jackson’s knowledge argument (1996, 103–104) in this connection. Finally, he suggests that there is available no analysis of the *concept* of phenomenal consciousness which might ground an entailment from the physical—basically, that functional and structural analyses of phenomenal consciousness are in principle inadequate (1996, 104–106, 111–121). Each of these five arguments operates by making vivid, in different ways, the fact that, even when we are being as careful as possible, we can imagine the physical remaining identical while the phenomenal varies.

Chalmers’ modal argument against materialism is thus at bottom an argument from conceivability, and all the force it has is whatever force an argument of this form can muster. However, because of the dilemma inherent in all arguments from conceivability, the fact that Chalmers believes he can coherently imagine a zombie world fails to show that zombie worlds really are possible, *even* if we grant Chalmers his claim (*pace* Kripke) that *a priori* ideal conceivability is enough to establish objective possibility. This is simply because it remains to be established whether it *is* in fact possible to fully understand all of physics, and to have complete knowledge of all actual physical states, and nevertheless to coherently imagine that conscious mental states are absent. The evidence and arguments which Chalmers gives in his book and elsewhere are sufficient to establish that we *think* we can ideally conceive of zombie worlds, but—without a second, lower level, modal argument, which he does not attempt to provide (and which would itself be inconclusive even if he did provide it)—they are insufficient to show that we really can ideally conceive of them. *This*, rather than his

modal rationalism or his arguments for *de facto* zombie conceivability, is the central problem with Chalmers' *a priori* attack on materialism.

To put the same objection in another way: we are not currently in a relevantly ideal epistemic situation with respect to the claims made by a completed, true physical theory. There is a vast amount of new information such a theory would provide that we presently lack, and this new information might well be relevant to the physical status of consciousness. Our present physical theories are almost certainly mistaken in some crucial respects, and are certainly incomplete, and thus we may currently be committed to distorting beliefs about physics. And it is not even clear—as Colin McGinn and others have argued—that we have the cognitive capacity to grasp the ‘final true theory’ of the natural world. The fact that we are not currently in relevantly ideal epistemic circumstances with respect to future science does not itself mean that we cannot ideally conceive of, or ideally fail to find imaginable, various non-actual situations—otherwise ideal conceivability would be a thin gruel indeed—but it does mean that *arguments* from conceivability in that domain must in principle be rationally unconvincing, for the reasons given above.

How might Chalmers respond to this? One route he might take would be to issue the following challenge. Since logical necessities are discoverable *a priori*, Chalmers might say, the challenge for his opponents is to find an *a priori proof* that zombie worlds are logically impossible, and the absence of such a proof is itself very strong evidence that none is available. In other words, if zombie worlds were impossible we should be able to prove it *a priori*—on the basis of the primary intensions of the concepts involved, as Chalmers would say. Since we cannot produce this proof, zombie worlds must be physically possible ... and, perhaps, it is the *non-availability of such an a priori proof* that Chalmers really finds so intuitively compelling (rather than merely the blank ‘imaginability’ of zombies).

But this cannot be quite right: for, notice the role that the term “*a priori*” is actually playing in Chalmers' account. His claim is not that we could have predicted the existence of qualia, or even the entailment of qualia by the physical, *in utero*, so to speak. His claim is that *given we*

know all there is to know about the physical fact-totality we could predict the existence and disposition of qualia. Similarly, Chalmers does not assert that we could *always have predicted* that life, for example, was physically reducible, just on the basis of a consideration of the concepts involved:¹⁴ instead, he says, once we found out enough about the relevant physical mechanisms we could *then* see that these mechanisms necessitated, and so explained, the phenomenon of life. Reducible properties are predictable *a priori* from sufficiently complete knowledge of their physical bases; without such knowledge, they need have no *a priori* connection to the physical at all.

The moral I draw from this is the following: an argument from conceivability of the type used by Chalmers *presupposes* that we have adequately complete knowledge of the micro-physical subvenient for conscious experience; however, there is currently no reason at all to think our knowledge of the physics of consciousness is complete, and attempts to conceive of possible worlds in which it is complete and in which we find zombie worlds conceivable are doomed to a vicious infinite regress of justification.¹⁵ Thus, Chalmers' argument is unsound.

Word count: 4,749

REFERENCES

- Chalmers, David. *The Conscious Mind*. Oxford: Oxford University Press, 1996.
- Chalmers, David. "Materialism and the Metaphysics of Modality." *Philosophy and Phenomenological Research* 54 (1999). 473–496.
- Chalmers, David. "Does Conceivability Entail Possibility?" In *Imagination, Conceivability, and Possibility*, edited by Tamar Gendler and John Hawthorne, Oxford: Oxford University Press, 2002. 145–200.
- Churchland, Patricia. "The Hornswoggle Problem." *Journal of Consciousness Studies* 3 (1996). 402–408.
- Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.
- Levine, Joseph. "Conceivability and the Metaphysics of Mind." *Noûs* 32 (1998). 449–480.
- MacPherson, Brian. "A Challenge to the Kripke/Putnam Distinction Between Epistemic and Metaphysical Necessity." *Southwest Philosophy Review* 13 (1997). 113–128.
- Yablo, Stephen. "Is Conceivability a Guide to Possibility?" *Philosophy and Phenomenological Research* 53 (1993). 1–42.

NOTES

¹ For yet another example, consider the free will debate: the fact that, in any normal situation, we can imagine ourselves having done something other than we did is sometimes taken as establishing that it was possible for us to have acted differently and that, therefore, our choice was free.

² This account of ideal conceivability is vastly less sophisticated and nuanced than those that appear in the articles mentioned above, but then it does not require their sophistication for the purposes of this paper. Its role here is merely to establish a lowest common denominator for accounts of ideal conceivability—to show that *some* account of epistemically ideal conditions is needed, even without saying precisely *what* account is best.

³ It can remain an open question for us whether the cognitive capacities of human beings are adequate for making modal judgements, bearing in mind that if they are not then modal scepticism looms.

⁴ I owe this example to Thomas Rauchenstein. This additional belief is distorting and false, but beliefs can be distorting even if they are true. Consider someone who believes that miracles sometimes occur which violate the laws of nature, and that therefore faster-than-light travel is possible in worlds nomologically identical with the actual world. Perhaps miracles do occur in some worlds nomologically identical with this one—i.e. perhaps the occasional violation of the laws of physics is consistent with the holding of those laws—but perhaps God has decreed that no miracles will actually occur, i.e. it is not possible for such violations to occur in the *actual* world.

⁵ That is, *q* is a modal claim—a claim about a non-actual possibility—as long as (as is surely very often the case) we are not *already* in relevantly ideal epistemic circumstances in the actual

world. More on this below.

⁶ [This, of course, is a crucial step in my argument, and it requires further consideration and defence. Chalmers, for example, suggests that “as long as one can know *some* claims about what is a priori or necessary with justification, an infinite regress is avoided, and I think pretty clearly one can (e.g. $1+1=2$)” (personal communication). My response to this is, roughly, going to be to agree that some modal claims—such as arithmetical necessities—are justifiable because we are sufficiently sure that we are *presently* in relevantly ideal epistemic circumstances, but that such modal premises can never be sufficient to establish modal claims with respect to which we are *not* presently in ideal epistemic circumstances. More pressingly, however, Chalmers also hints that there might be *other* ways of arguing for contested modal conclusions than by appeal to arguments from conceivability: he suggests using “general first-order reasoning using our concepts, or some sort of conceptual analysis, rather than conceivability arguments per se” (personal communication). I get into this a little below, but I think more will have to be said in future iterations of this paper: in general, it seems that this pushes Chalmers into placing the whole weight of his modal argument on the claim that it is an *analytic* truth that, say, zombies are not logically possible—and this would be, perhaps, a) interesting in itself, and b) maybe too implausible a claim to be supported. I would welcome input on this issue from my commentator.]

⁷ It is somewhat more plausible to suppose that, for some reason, our judgements of conceivability *when we are in epistemically ideal conditions* reliably track logical possibility. But that kind of case is not what is at issue here, since we are presently dealing with modal claims that are made in non-ideal epistemic circumstances (in just the same way as most knowledge claims are made in non-ideal epistemic circumstances).

⁸ This argument against arguments from conceivability resembles Hume's attack on induction in the following way: just as Hume argued neither that inductive claims are false, nor that we should not (defeasibly) make inductive judgements, the argument presented here against arguments from conceivability does not render worthless the currency of modal intuition. I do not rule out—and in fact am concerned to hold open—the practice of making modal judgements on the basis of conceivability, and of sticking to such judgements unless some new evidence comes to light (new background information or the removal of distorting beliefs, for example) which might overrule them. What Hume and I both dispute, however, is the *rational justification* of judgements in these respective spheres.

⁹ Or at least, the move from ideal conceivability to the modal sub-conclusion within such arguments.

¹⁰ Note that this is not to say that obvious modal claims cannot be properly used in interesting arguments; in such cases, the modal claim involved will only be 'uninteresting' *qua* modal claim, rather than *simpliciter*. For example, Kripke uses a range of modal claims in "Naming and Necessity" which—though in themselves 'obvious' (or at least, Kripke does not seem to think they require any defence)—have far from obvious metaphysical consequences. I am indebted to William Seager for this point.

¹¹ The restriction to "the *positive* facts about consciousness" is intended to deal with the problem, in defining physicalism, of the possibility of there being *extra* unphysical stuff. Chalmers' solution, in brief, is to formulate the supervenience claim as dealing only with the set of supervenient facts true of *this* world, and then restricts this set of supervenient facts still further to include only "positive facts," i.e. those which do not include or entail negative existential claims. It would be a problem for Chalmers' unmodified world-specific

formulation if fixing the physical facts also fixed such facts as that there are no angels or that all living things are based upon DNA—roughly, this is problematic because it apparently makes physical claims like these *a priori* rather than empirical.

¹² Chalmers distinguishes between what he calls type-A and type-B materialism. Type-A materialists hold that “phenomenal truths are necessitated *a priori* by physical truths” (Chalmers 1999, 473); that is, for our purposes, they hold that materialism is inconsistent with logically possible zombie worlds. Type-B materialists “accept that phenomenal truths are not necessitated *a priori* by physical truths, but hold that they are necessitated *a posteriori* by physical truths” (Chalmers 1999, 474); we can say that these materialists take their position to be falsified only by metaphysical, and not purely logical, possibilities. Chalmers argues that type-B materialists are *wrong* to think that the ideal conceivability of zombie worlds is compatible with materialism.

¹³ Interesting difficulties arise when we try and explicate logical necessity—or what Putnam sometimes calls “epistemic necessity”—using the language of possible worlds, which Putnam and Kripke do not attempt to do in any extensive way; see MacPherson 1997 for a discussion of this point. Be that as it may, however, another way of putting this view, presumably, is to say that “the maximal set of possible worlds is the set of metaphysically possible worlds, and when we say that some situation is possible, without qualification, we mean that there is a metaphysically possible world in which it obtains” (Levine 1998, 450). Thus, whether or not something is logically possible, if it is metaphysically impossible then it is simply objectively impossible.

¹⁴ “Vitalism was mostly driven by doubt about whether physical mechanisms could perform all the complex *functions* associated with life: adaptive behavior, reproduction, and the like”

(1996, 109).

¹⁵ As Patricia Churchland puts it: “Perhaps we like to put our ignorance in a positive light, supposing that, but for the Profundity of the phenomenon, we *would* have knowledge. But there are many reasons for not knowing, and the specialness of the phenomenon is, quite regularly, not the real reason.” (1996, 406)