

PAUL BLOOMFIELD

LET'S BE REALISTIC ABOUT SERIOUS METAPHYSICS¹

This is a warning against arguing directly from the logical possibility of concepts to the real possibility of things. Immanuel Kant

In the history of western philosophy, we find two traditional views of “serious metaphysics”: Aristotle’s and Kant’s. Albeit not as clear as we might prefer, from Aristotle we learn that metaphysics is the study of “being *qua* being”. We take this to mean a study of reality as it is, in itself, independent of how we conceptualize it. From Kant, we learn that this is not quite possible, that our conceptual apparatus is employed in all thought, even thought about “being *qua* being”. For him, knowledge of metaphysics is encompassed by necessary knowledge that is synthetic *a priori*, derived by transcendental argument moving from our experiences to their necessary prerequisites. While neither put it perhaps in exactly these words, both Aristotle and Kant took their subject matter to be the actual world as it is, and not merely the study of a world as it could possibly be but actually is not. On one hand, considering how things possibly are is useful to metaphysics insofar as this shines light on how things actually are. On the other hand, however, from considering possibility for its own sake, or possibility not predicated upon that which is actual, we gain no insight into the actual world and we are not doing a good job with our metaphysics. Undoubtedly, our technical understanding of modality has become more nuanced than it was for these old world philosophers. Nevertheless, they have determined the parameters of the subject matter for serious metaphysicians.

As in introduction to what is meant here by “serious metaphysics”, Kant’s second analogy can serve as a paradigm: we find it a metaphysically necessary, synthetic *a priori* truth that causes temporally precede their effects. This is not a causal truth, but rather a condition on the application of the concept of causation. Neither is this a truth of logic or a conceptual truth because one cannot derive a contradiction from the claim that it is false. (This is saying nothing more than that it is a synthetic and not an analytic truth.²) There is no logical contradiction in considering a world where effects precede their causes, as there is in considering a world with

square circles. There is, then, a sense of modality in which it is proper to say that a world with backwards causation is logically or conceptually possible, but we must also note that this possibility is irrelevant from the point of view of the metaphysics of the actual world because it (the actual world) cannot be as this possibility allows. There is a logically possible world at which there is backwards causation, nevertheless it is metaphysically impossible at the actual world. Merely being “logically possible” does not make something metaphysically significant and those doing serious work on the metaphysics of causation need not, indeed ought not to consider this logical possibility when constructing their positive theory of the nature of causation. This possibility should not be allowed in any serious metaphysics of causation. If instead we consider what is actually true as constraining the sense of modality being employed here, and ask what is necessary or possible *assuming or given the actual world as it is in itself*, then we are engaged a serious metaphysics. In other words, metaphysicians should not to begin by asking “what are the possible ways the actual world could be?”, as this employs a fully unconstrained notion of “possibility”: it is not constrained by how the actual world actually is. Rather, metaphysicians should begin with the actual world as it actually is and then ask what is possible given this constraint of how actual world actually is.

Another way to explain serious metaphysics is to note that serious metaphysicians think there is something ontologically special about our world insofar as it is the only world that can properly be considered as actual. Leaving Lewis’ (1986) modal realism out of the picture for the moment (as he thinks other possible worlds exist in the same way as the actual world), “actuality” is not merely an indexical; it cannot be shifted from world to world; the only world fit to be considered “as actual” is ours. In this sense, serious metaphysicians use an elitist’s asymmetric modality. The contents of our actual world are ontologically different than the contents of any other possible world. Philosophers can treat a world with unicorns “as actual” and ask what else would be possible if this world were actual, but serious metaphysicians should not such questions. Instead, serious metaphysicians should only be interested in possibility insofar it sheds light on the metaphysics of the actual world. Perhaps what is special about the actual world is that it is merely where we are, but serious metaphysicians should think that the single actual world is metaphysically of a different nature than all the other possible worlds.

Leibniz was not, for example, arguing that it is possibly true that we are monads; rather he was arguing that this is actually true. He was not suggesting that God could have created the world in which we are monads

but in fact did not create that world. Insofar as Berkeley denied the claims of “the metaphysicians”, he was *ipso facto* doing metaphysics himself and he was not arguing that it is possibly true that there is no matter (or that this possibility is logically consistent with what we actually observe), but was instead arguing that there is no matter in the actual world. Hume, also fond of using “metaphysics” pejoratively, was nevertheless a serious metaphysician: he did not argue that it was possible that he had no self but that he actually had no self. (Note Hume’s theory of causation does not accommodate the possibility of backwards causation, even though this is, in some sense, “possible”.) Serious metaphysicians take their subject matter to be the metaphysics of the actual world and the class or set of worlds which are possible given the stipulation that the actual world is actual; their discussions are always grounded in, or indexed to, what is real, what is actual, and not in what is merely logically possible or what fails to logically contradict what we know. The metaphysics of those worlds which could not be possible, when the actual world constrains the sense of what is “possible”, are nothing more than red herring.

The constrained sense of “possibility” referred to above has received a large amount of attention in the literature; in particular, its technical aspects were first introduced by Robert Stalnaker’s “Assertion” in 1978, and further explicated by Martin Davies and Lloyd Humberstone in their paper “Two Notions of Necessity” (1980). Here they lay out both an unconstrained and a constrained sense of modality. In the unconstrained sense, actuality is treated as a mere indexical, such that counterfactual worlds may be considered actual. This is the sense of “possibility” indicated above, wherein we considered what would be possible if there were actually unicorns. This sense of “possibility” is given by logical consistency: if a world is logically consistent then it is possible in this sense; we ask what is logically consistent with the “actual” existence of unicorns. The constrained, indexed, or centered sense of “possibility” is one which, as mentioned above, takes the actual world as a given and asks what is possible given this. Using this sense of “possibility”, we ignore what is possible given the “actual” existence of unicorns because unicorns do not actually exist. With this centered notion of modality, we begin with the actual world and consider what is possible, consider that set of possible worlds, based on this. The thesis to be explored below is that serious metaphysicians should be concerned with the latter centered, as opposed to the former uncentered, sense of “possibility”: we should continue on the path begun by Aristotle and followed by serious metaphysicians since then, or up until recently. Serious metaphysicians should not be asking “what is possible?” *simpli-citer*, but rather “given what is actual, what is possible?”; we should be

concerned with the actual world and the set of worlds which are possible, given the centered sense of “possibility”.

For philosophers have lately seen this old school primacy given to actual reality demoted to secondary importance, while the primary focus of metaphysics has been switched to what is merely logically or conceptually possible, using the unconstrained notion of “possibility”. This new trend in metaphysical methodology comes to us from, among others, the work of David Chalmers (1996) and Frank Jackson (1998), and in a joint paper of their’s from 2001. These philosophers hold that we can gain *a priori* knowledge of metaphysics by attending solely to our concepts when these are considered independent of how things actually are. I call this a “method” because it is the practical application to metaphysics of work on the formal, logical systems of modality in which “actuality” is treated as an indexical. I think the application is misbegotten because it garners putative “knowledge” about the nature of reality, our actual reality, by attending to worlds which are logically consistent with what we observe at our world but which are nevertheless impossible *given what is actually true*. Attending to these actually impossible worlds yields only metaphysical “knowledge” which is actually founded on ignorance.³ These philosophers take worlds which could not be actual but are still “possible” (in the sense of not being logically contradictory) to be nevertheless helpful to us in learning about the metaphysics of the actual world. This is a mistake.

To be quite clear: the criticisms below do not concern how the method to be discussed is employed or the particular entailment relations which the method condones; rather, the problem is with the method itself. The criticism does not concern how the philosophical machinery is used, but rather concerns how the machine is constructed. For example, I will not accuse Chalmers or Jackson of making an invalid inference in the derivation of any entailment, but will accuse them of using a faulty formal system which gives license to certain inferences as “valid” (given that system) which should not be considered valid by serious metaphysicians. One might say that the problems to be discussed are not philosophical, but meta-philosophical: this is an essay in meta-metaphysics.

1. THE CJ VIEW

I call the Chalmers/Jackson method the “CJ method” or the “CJ view”, and stick for the most part to Chalmers’ version of the jargon. Focusing on the example of the meaning of “gold” is apt, first because Saul Kripke’s discussion (1972) of gold and fool’s gold is so helpful, and second in order to introduce John Locke’s salutary discussion (1975 [1689]) of nominal

and real essences to these matters. I conclude with two sections addressing possible defenses of both Chalmers and Jackson in turn, and then a section on epistemic possibility as discussed in Chalmers and Jackson (2001) and in Chalmers (as posted on the internet, see Bibliography at end).

This is how the CJ view is supposed to work. We begin by noting that “gold” has two meanings. The first is that “gold” is the name of a complex concept or an idea of “a Body yellow, of a certain weight, malleable, fusible, and fixed” (Locke 1975, 439). It is supposed that this meaning of “gold” (when apt, referred to as “p-gold” below) derives from the fact that while we have causal contact with gold, initially we are ignorant of gold’s underlying nature or what makes gold be gold; for want of knowledge of the “deep” nature (metaphysical nature) of gold, we use the superficial characteristics of gold, like its color and malleability (etc.), to later reidentify gold when it passes our way again. According to the CJ view, this meaning of “gold” deserves the appellation of the “primary meaning” of the term. There is, however, another meaning to “gold”, a “secondary meaning”, sometimes below called “s-gold”, in which the word “[s-]gold” refers directly to the underlying constitution of the stuff that chemists have more recently taught us is the element with atomic number 79, abbreviated as “Au”.

Assuming the truth of chemistry and given this secondary sense of “gold” alone, being s-gold is necessarily identical with being Au, according to the CJ view. If, however, we confine our attention to the primary sense of “gold”, p-gold, it is possible for p-gold not to be Au: for all we know (note the epistemic modality), according the CJ view, it could have turned out that the yellow, malleable stuff which we call “gold” in the primary sense had had the underlying nature of something other than Au. And we can know this about p-gold *a priori*: we need not know anything about the nature of the empirical or actual world to know that p-gold could have turned out to be something other than s-gold; indeed, before the advent of chemistry, we had no clue at all as to what s-gold was. So, there is no logical contradiction in the idea of p-gold which is not s-gold. Thus, a good reason to think the so-called primary sense of “gold” really is primary is because we could use the word “(p-)gold” in all sorts of interesting and useful ways without knowing anything about what we are here calling the “secondary meaning” of “gold”, namely that s-gold is identical to Au. What fixes the primary meaning of “gold” is not what fixes its secondary meaning. Its secondary meaning is determined by the metaphysical nature of the actual stuff baptized “gold” or what Locke would have called an “I know not what” and what we now call “Au”. The primary meaning of “gold” is determined by our perceptions of s-gold. *Contra* the CJ view,

this is a good reason for thinking that the “secondary” meaning of “gold” is metaphysically primary: even though we learn the meaning of “s-gold” at a later time than we learn the meaning of “p-gold”, s-gold is metaphysically primary. The metaphysics of the stuff we call “gold” are to be found in its secondary sense.

If the meaning of “p-gold” is, roughly, a list of superficial characteristics and if s-gold is only the stuff at the actual world which possess those superficial characteristics, then everything we actually call “gold” will be what we think it is. There will be no mistake; our ignorance of the meaning of “s-gold” will have no effects. If, however, there is some sort of “fool’s gold” in the actual world, then we may sometimes be calling something “gold” when it actually is not. “P-gold” does not discriminate between s-gold and fool’s gold; we may unwittingly make mistakes in calling things “gold”. If we try to draw metaphysical conclusions about gold based on the primary meaning of “gold”, then we may very well be drawing false conclusions. We may derive “facts” about gold from the primary meaning of “gold” which are not actually true of s-gold. The world logically could have been such that these conclusions are true, but in fact they may not be true. We let our ignorance of the actual world play an unnecessary role in our thought if we attend to primary meanings without regard to secondary meanings. If we try to draw conclusions about the metaphysics of gold based solely on the primary meaning of “gold” then we may be led astray: there may be metaphysical stories one could tell about gold which are logically consistent with the primary meaning of “gold” but which are nevertheless not possible given what is true at the actual world. It is logically consistent with the primary meaning of “gold” that the stuff we call “gold” is really fool’s gold or perhaps a psychic substance which deludes us into thinking it is yellow, malleable, etc. This however is actually impossible *given the way the world actually is*. If the world actually is as it is (and how could it fail to be so?) and if we pay no heed at all to this and consider in our metaphysical theorizing logically possible worlds which are not possible given the actual world, we allow in to our theorizing a greater chance of error than required: we consider certain possibilities as metaphysically significant due to our ignorance of how things really are. We speculate without keeping an eye on how the world actually is and this is “metaphysics” in the properly pejorative sense. The consideration of such possibilities warrants the consideration of what is not possible at the actual world and this unnecessarily increases the chance that we will be lead astray from the actual truth regarding this world’s metaphysical structure.

To see the philosophical force of this way of thinking of things, we need only to attend to Chalmers' argument for dualism. In the briefest of terms, we find that "consciousness" – in the primary sense of "consciousness" – need not be attached at all to brains or to anything material. In the same way that it is possible for all things p-gold to not be identical to Au (since it may be fool's gold), it is possible for creatures, zombies, to behave and physically be in every way just like us and still not be conscious. The idea of a zombie is not self-contradictory. Zombies are not logically or conceptually impossible just as p-gold that fails to be Au is not impossible. If so, then consciousness need not be identical to something in a physical duplicate of one of us and this failure of identity entails a form of metaphysical dualism.

The CJ method, as applied to consciousness, may be put in a more colloquial formulation taken from Kripke (1972, 153ff), to which we will return below: God had more work to do to make us conscious after he made us with the brains we've got and we know this, *a priori*, because of the possibility of zombies. Fixing our physical nature does not logically or conceptually guarantee our consciousness because zombies are *in this sense* possible. Even though, at the actual world, the mental causally (naturally) supervenes on the physical and they cannot be pried apart causally here, this does not imply that they cannot be logically pried apart. As noted, the concept of a zombie is not self-contradictory and is therefore logically "possible" in this sense. To return to our theodicy, when God was choosing which world to create he could not simply choose one member of a set of worlds where creatures are physically constituted as we actually are, because logic tells us that there are some worlds in this set which have creatures physically like us but without mental states. Therefore, our actual material nature does not logically entail our being conscious, and some sort of dualistic metaphysics of the mind becomes true by default. Dualism is in this sense necessarily true: at *any* world with sentient creatures, there will be still other possible worlds which have creatures just like those but without mental states. At every world with embodied sentience, dualism is (in this sense) possible and therefore true.

This method of doing serious metaphysics may be called into question by turning to John Locke for help: as both a traditional metaphysician and a scrupulous and modest empiricist, he is just the sort of philosopher we need at this point. In particular, we may appeal to his distinction between nominal and real essence (1975 [1689], book III, chap. VI).⁴ While it is not formulated in explicitly modal terms (though "essence" is implicitly modal), this is the same distinction – I submit – as the primary/secondary meaning distinction of the CJ view, though cast in a different, more per-

spicuous, light. In Locke, we find nominal essences play the same establishing role in naming a natural kind as primary meanings are said to play on the CJ view. Nominal essences are the meanings of these terms such that they are the collection of ideas or properties by which we name and reidentify tokens of species of things in the world: the nominal essence of “gold” is as quoted above, its yellowness, malleability, etc. Real essences are, on the other hand, the underlying “constitution of the insensible parts of that Body, on which those Qualities, and all other Properties of [things like] Gold depend” (p. 439).

So, at one level, Locke is in agreement with the CJ view: the meaning of “gold” for Locke is given by its nominal essence. On the other hand, he would not agree that we can gain (even defeasible) *a priori* knowledge of things in the world, we cannot do serious metaphysics, by analyzing (or finding out what is logically consistent with or entailed by) nominal essences, for these map only contingently onto real essences or the stuff of the actual world. In reference to “Sulfur, Antimony, or Vitriol” Locke says the following: “For though they are Bodies of the same *Species*, having the same nominal *Essence*, under the same Name; yet do they often, upon severe ways of examination, betray Qualities so different one from another as to frustrate the Expectation and Labor of very wary Chymists” (p. 443). It would be preferable to name things by their real essences, for were we to do so, all members of a species would necessarily exhibit the same properties and would not so frustrate the “Labor of [these] very wary Chymists”. Unfortunately, according to Locke, it is not possible to name things in this fashion, for in his characteristically modest way, he claims that as far as real essences are concerned “we know them not” for reasons having to do with what he took to be the limits of our epistemic abilities.

For Locke, the metaphysics of things, their primary essential nature in reality, is fixed by what the CJ view calls “secondary meanings”. Indeed, from the point of view of serious metaphysics and not the philosophy of language, we can find the prioritizing of Locke and the CJ view quite the opposite of one another: if we were not epistemically bound as we actually are, metaphysics would proceed first off, or primarily, by studying the real essences of things and leaving their nominal essences, their “primary meanings”, behind as secondary (and probably otiose).⁵ For Locke, this would be serious metaphysics.

So, where has the CJ view gone wrong? A fuller understanding may be obtained by re-introducing the modal element into the dialectic; and here I follow the seminal discussions of Stalnaker (1978), and Davies and Humberstone (1980). As Chalmers and Jackson recognize, the primary meanings of the CJ view have a modal import which is not indexed to the

actual world. The primary meaning of “gold” is one such that it is disattached from actuality or is an “uncentered” sense of possibility: worlds get hypothetically treated as actual which are not actual but counterfactual. Both primary and secondary meanings have implicit modal operators built into them. The former contains an uncentered sense of “possibility” in which all possible worlds may be considered as actual: all possible worlds are to this degree possible. The latter contains a centered sense of “possibility”: it assumes that the actual world is the only actual world and asks what is possible given this. With regard Stalnaker’s famous “diagonal proposition”, we note that actuality in it seamlessly slides from one world to the next, and in doing so we are asked to treat non-actual worlds as actual.

Hence, whatever is yellow, malleable, fusible and fixed is in this way independent of the real essence of gold at the actual world, because it is possible for the primary meaning of “gold” to refer to stuff which is not actually gold. It is not necessarily the case that all possible substances falling under the nominal essence of “gold” are examples of gold at the actual world. On the CJ view, what we call “gold” in the primary sense is only contingently “gold” in its secondary sense. But, as Putnam notes in “The Meaning of ‘Meaning’ ” (1975), while it is conceivable that what we call “gold” is not gold, it is (of course) not logically possible for gold to fail to be gold.⁶ Rather, what we should say is that we use nominal essences or primary meanings to help us pick out natural kinds, and the kinds are intended to be “natural” insofar as we intend natural kind terms, like “gold”, to pick out the real essence of those natural kinds or their term’s secondary meanings. It is only by distinguishing between nominal and real essence (primary and secondary meaning) that we can conceptually distinguish gold from fool’s gold. If we ignore the secondary meaning of “gold” and pay attention to only its primary meaning, we lose the ability to distinguish gold from fool’s gold and thereby the ability to learn about the actual nature of real gold. Zombies are the fool’s gold of consciousness studies.

The proponent of the CJ view may be impatient by this point to say something like the following:

Look, there is no possible world which has Au but no gold, while there is a possible zombie world which has brains with no consciousness. This shows us that being gold has a different metaphysical status than being conscious. So, the whole model you’ve created here has some sort of problem built into it.

Call this “the Response”.

Of course, the Response is understandable, but it misses the point. The possible world with zombies that the CJ view points to employs an uncentered sense of “possibility”; it is merely logically possible and of no use to us in determining the actual metaphysics of consciousness. As a response to the Response, this might sound question-begging, but the point can be put more neutrally in terms of the following dilemma: Take the real essence of consciousness, the secondary meaning of “consciousness”, whatever it may be, and call it “Bob”. Then call the following question “the Question”: Is there a possible world at which you can have Bob and no consciousness? Well, this depends on how the word “consciousness” is taken in the Question. If we take it to be the secondary meaning or real essence, then the answer is “No”, for the same reasons that there is no world with Au and no gold. The real essence fixes the metaphysics. The sense of “possibility” which goes beyond this, the one where possibility is constrained by only primary meanings and what does not logically contradict these, is not relevant to the more tightly constrained metaphysical project regarding real essences at the actual world. If in the Question we are talking about the primary meaning of “consciousness”, then the answer is “Yes”, because there is a logically possible world with all the superficial characteristics of consciousness but without Bob. Taking this seriously, however, is like consenting to the possibility of there being gold without Au and, while this may not be a logical contradiction, it is irrelevant from the point of view of learning the truth about metaphysics at the actual world. (The metaphysics of the actual world concerns what is possible given the actual world, and from this perspective it is not possible for there to be gold that is not Au; *mutatis mutandis*, the same can be said of consciousness.) Either way we take the Question, the CJ method leads us astray. Only ignorance of what is actually true could lead us to consider brains (like ours) without consciousness or gold without Au. Since these are actually impossible, they are metaphysically irrelevant.

Consider once again the first sentence of the Response. When inspected very carefully, one can detect a subtle slide from the use of secondary meaning to primary meaning. In the first part of the sentence about gold, it must be the secondary meaning of “gold” which is being employed, because if it were the primary meaning we could point to a cold world with odd light where Au was neither malleable nor yellow. In the second part of this sentence, it must be the primary meaning of “consciousness” being employed because if it were the secondary meaning we would have a case where consciousness comes apart from its real essence and that would be like saying there was a world with Au and no gold. This cannot be.

The upshot is that when doing serious metaphysics, we should attend first and foremost to what the CJ view deems “secondary”. What is possible will give us insight into actual reality only if the possibilities we are considering are truly possible for us at the actual world. This requires a centered sense of modality be employed while investigating the metaphysics of the actual world. We are left to consider in our metaphysical investigations only those possibilities that remain possible after we have indexed our modality to the actual world. We should ignore the uncentered possibilities attached to the CJ view’s primary meanings. These uncentered possibilities, it might be said, are “possible” in the sense given above but are not, one might say, “actually possible”. It would, I should think, be pointless to assert that there is no sense of “possibility” which names these uncentered possibilities and indeed, as we will see below, the best sense to make out of this notion of “possibility” is to consider it as an epistemic modality. Serious metaphysics should not be concerned with all possible knowledge, however. It is concerned with knowledge about the actual world and its metaphysical basis, which cannot be studied without considering what is possible *given the actual world*. We are not interested in the metaphysics of those worlds which are impossible given the actual world. Indeed, when we are thinking clearly, when we are doing our best to avoid mere fancy or foolish folly (or, *pace* Lewis, incredulous stares), we think there is only one actual world, one reality, and we – you and I, reader and author – are living in it.

It seems to be a fact that things could actually be different than they are, and this fact is important to us and to serious metaphysics insofar the possible worlds which represent these different “ways things could be” can help us learn about our actual world. We think it is important that Humphrey could have won the 1968 U.S. election, for this possibility yields insights into what we (conscious) human beings can actually do. The present criticism of the CJ view as a method of serious metaphysics parallels, in one way, Kripke’s (1972) criticism of Lewis’ (1968) counterpart theory: Kripke says we are not concerned with what is the case for our counterparts, regardless of how *similar* to us they are; rather, we are concerned with what is possible for us here at the actual world.⁷ The problem with the CJ method for doing “serious metaphysics” is that it tells us to be concerned with something which is disengaged from the actual world or reality, and so leads us away from the explanandum of serious metaphysics. Zombies are the creatures of confusion, even though they are not self-contradictory, and allowing them into the discussion only breeds further confusion. Though it is not a logical contradiction to think that you or I could have been born of different parents, or that effects could

have preceded causes, these possibilities are otiose to our understanding of the way things actually are and to metaphysics when this is taken as the study of actual reality. Given the actual world and its facts about us as *Homo Sapiens*, we do not think it is not possible for one of us to be born of different parents and we ought not to consider this possibility when seriously doing the metaphysics of personal identity. For perfectly analogous reasons, we ought not to consider the possibility of zombies when doing the metaphysics of consciousness.

To conclude the main portion of the paper, here is a plausible etiological diagnosis for how the CJ view goes wrong: proponents of the view have been lured in by the egalitarian elegance of formal modal systems wherein all possible worlds stand on all-fours with respect to each other, such that actuality can be transferred to any world by means of a formal notational device working like an indexical. "Actuality", on this view, can switch reference just as the reference of the word "here" switches. As noted above, this is most especially true given Lewis' modal realism where all possible worlds are equally actual if each is taken from its own point of view. But if we are not modal realists then our actual world is existentially unique; it is the only real, existing world. If we take possible worlds to be (say) sets of propositions, no one should then go on to mistake any set of propositions for the actual world, even though the actual world is also possible. (There is a set of propositions which represents the actual world.) The meaning of "actual" is different in those uses where it refers to the actual world than it is in counterfactual uses, when other possible worlds are considered as "actual"; the formal point of the argument above is to take note of this equivocation and to point out its relevance to "serious metaphysics". But even if we are not modal realists, we may still use counterfactual language which leads us to speak as if other possible worlds can be treated as "actual". Speaking this way insidiously leads us to understand the difference between the actual world and all the other non-actual worlds as a purely logical or formal difference. One way to make this explicit is to employ a double-indexing device. We can say that, from the CJ point of view, the world which is actually actual (namely our world) has no more theoretical or metaphysical significance than any other world, all of which are possibly actual.

Metaphysicians have gotten used to thinking in terms of having to ask "which world are we now taking as actual?", when in fact this question is itself a manifestation of metaphysical confusion. In doing serious metaphysics, we never have to ask which world is properly taken to be actual, the answer to that is obvious and never changes: actuality does not ever move or slide from this, our actually actual world. The difference between

our actual actually world and all the possibly actual worlds is metaphysical, ontological, and not purely logical (much less conceptual). The actual world is metaphysically unique and serious metaphysics takes it as its subject matter.

2. TWO DEFENSES OF CHALMERS

The most likely place for a defender of Chalmers to object is in regard to the relevance of zombies to the metaphysics of consciousness. This would come by way of an argument concluding that, with regard to “consciousness”, its primary and secondary meanings are necessarily the same. Thus, they cannot come apart as is the case with “gold”. In Lockean terms, the real essence of pain is, on this view, necessarily the nominal essence of pain. So, the dilemma presented above with regard to the underlying nature of consciousness (there called “Bob”) being present without consciousness being present is seen as malformed from the start.

The response to this is unfortunately dialectical, but follows directly from the type of criticism presently being leveled at the CJ view. This is to say that the arguments above were aimed at the methodological machinery which is only then later used to establish the conclusion that the primary and secondary meaning of “consciousness” cannot come apart. The (hard?) problem of consciousness should not be located within the relationship between the meanings of the word “consciousness” and consciousness, rather the problem is (within) the actual nature of consciousness itself alone.

This response is apt because it is Chalmers who brought zombies into the picture in the first place, and so he is the person who (1) insists on a possible gap between having a brain like ours and being conscious. Later however, it turns out that (2) there can be no logical gap between the underlying nature of consciousness and being conscious. And so some sort of dualism must be true. The point of disagreement is not with the move from (1) to (2), as the defender of Chalmers claims. Rather, the problem arises before this, with the logical machinery that brings (1) itself into the dialectic as relevant, the point at which the logical possibility of a zombie is introduced even though zombies are actually impossible. The ontological importance of (2) does not get off the ground without (1) and this is where the problem with the CJ view resides.

Perhaps the defender of Chalmers will insist something like the following:⁸

What we (Kripke included) should say is that all God needed to do to give us consciousness was give us brains and bodies like we've got, and if he wanted to make a zombie, he would have to make the further move of breaking the actual laws of nature. It does not take a miracle at the actual world to get consciousness out of the brains we've got, but it would take a miracle to make a zombie at the actual world.

Chalmers could agree to this, saying that the mental naturally supervenes on the physical, and so he does not think that mental states and brain states can come apart from physical states at the actual world. Given that some legitimate sense can be made of primary meanings and that doing conceptual analysis on them does not lead us to think that zombies are possible at the actual world (something to which all parties agree), where exactly is the mistake of the CJ view? Again, the answer is that problem is not with how Chalmers employs his machinery, there is no mistake at this level, rather the problem is with the machinery itself.

The point to question Chalmers at comes after he lays out the distinction between primary and secondary meanings and before he goes onto employ it. At this point, we must ask what pre-theoretical reasons there are to think that we can find out about the nature of consciousness based on the primary meaning of "consciousness". The answer is that we have intuitions about zombies, but this is no answer in the face of their actual impossibility and this is most especially so when we remember that we form primary meanings in ignorance of the actual facts. We do not claim to get any knowledge of horses by thinking about Pegasus, and for the same reasons we should not claim to get any knowledge about consciousness by thinking about zombies. Chalmers is arguing that, at any world where material creatures are conscious, those mental properties do not logically supervene on the physical properties of those creatures, thus dualism is necessarily true (which includes being true at the actual world). The argument is based on the logical possibility of zombies, or the fact that the concept of a zombie is not self-contradictory. If Anslem's Ontological Argument was criticized for trying to prove a truth about God by the inspection of our concept of "God", Chalmers is open to the same sort of criticism. Indeed, one could quote from Kant's First Critique response to the ontological argument as a response to Chalmers: "There is already a contradiction in introducing the concept of existence – no matter under what title it may be disguised – into the concept of a thing [in this case, a zombie] which we profess to be thinking solely in reference to its possibility" (A597;B625). The bare logical possibility of a "zombie" is ontologically vacuous.

The final challenge to Chalmers is a footnote from the same page of the *Critique of Pure Reason*:

A concept is always possible if it is not self-contradictory. This is the logical criterion of possibility, and by it the object of the concept is distinguishable from the *nihil negativum*. But it may none the less be an empty concept, unless the objective reality of the synthesis through which the concept is generated has been specifically proven; and such proof, as we have shown above, rests on principles of possible experience, and not on the principle of analysis (the law of contradiction). This is a warning against arguing directly from the logical possibility of concepts to the real possibility of things.

3. SOME DEFENSES OF JACKSON

Jackson does deliberately deal with many of the issues addressed above in his defense of conceptual analysis as properly having a role in doing serious metaphysics. In particular, Jackson makes a distinction between modest and immodest uses of conceptual analysis and claims that the modest use to which he limits himself does not play “a major role in an argument concerning what the world is like” (p. 43). Conceptual analysis begins with the concepts of folk theory, but Jackson recognizes that these are fallible; they are not “sacrosanct”. We may learn, for instance, that there is nothing in the world which corresponds to our folk concept of “free will” and so we may replace the folk concept with another which answers adequately to our original need for a concept of “free will” but does not come with the ontological baggage that makes it incompatible with determinism. So, there is no argument from how our concepts are to how the world must be. We must be open to the possibility that our concepts are in error, and the world helps us to see these errors.

The first response is to note that whatever “serious metaphysics” might mean, it has to have something to do with giving arguments that concern what “the world is like”. If conceptual analysis does not play a major role in this project, then it plays no major role in serious metaphysics either. If the world is to show us that our concepts are error-ridden, then the serious metaphysics happens when we find out how the world actually is. And if we do, in the end (and despite any protest), allow a conceptual analysis of some suitably modified folk theory, in the form of primary/A meanings, to play some role in determining what we are looking for out there in the world, serious metaphysics is only this latter project of “looking” and will begin only after the conceptual analysis is finished. Even if we countenanced the conceptual analysis of primary intensions, the task of finding out what exists only would begin after we let our conceptual analysis “fix” what it is we are looking for. *Whatever it is* that fixes what

we are looking for, maybe the world contains it and maybe the world does not; it is only in investigating this latter point that marks the *beginning* of serious metaphysics.

In any case, it is not obvious that Jackson has kept his conceptual analysis as modest as he claims. He does end up with some very powerful results (which is what makes his book so excellent), and it would be hard to deny that he relies heavily on primary/A meanings or intensions in getting them. The problem, as pointed out above, is that however well and suitably we may modify the primary/A meanings, they are forever and only contingently related to what exists in the world. To find out what is in the world requires us to go primarily to the world for our data, not our concepts. The actual world should be kept at the forefront of our minds, for we are only concerned with what is possible when it is assumed that what is actual is a given. We should keep our modal intuitions narrowed to those indexed to the actual world: secondary/C extensions and meanings. Data about how we talk or think, data about primary/A meanings *might* help to confirm a theory, but it certainly ought not to set the explanandum of the theory.

A defender of Jackson will not likely be impressed by this response and indeed more must be said. We are asked by Jackson to consider the example of the analysis of “knowledge” as “justified true belief” and to the role that Gettier (1963) gave conceptual analysis in showing us that this understanding of knowledge was incomplete. Here the situation is different than in the case of “free will” for Gettier explicitly appealed to possible cases in order to show that knowledge was something more than justified true belief. Jackson takes it that his method of serious metaphysics (using primary/A intensions) improves on traditional conceptual analysis in the way that Gettier’s work was an improvement on Ayer’s and Chisholm’s conceptual analysis (see 1998, pp. 28–29, 47). (The difference between the Gettier and the free will case is that with free will it is the scientific or empirical “facts” of determinism that move us to modify our concept; with Gettier, we are moved by pure philosophy.)

There is no doubt that Gettier’s counterexamples serve as a paradigm example of a successful move in analytic philosophy, but this does not really help the CJ view. Gettier’s examples were so moving because they were *actually possible*. There are possible situations in the actual world where a person may be said to have true justified beliefs and yet intuitively may not be said to know. Gettier is pointing to modalities which are indexed to actuality in order to make his point; if had pointed to an *actually impossible* situation, or to a situation which was merely not self-contradictory but not actually possible either, in which justified true be-

belief was shown to co-exist with a lack of knowledge, he would not have changed as many minds as he did. Though a bit anachronistic, we may say that Gettier was concerned with the secondary meanings of “justified true belief” and “knowledge” and not with their primary meanings.

Jackson thinks that his use of conceptual analysis is modest because he does not treat concepts as sacrosanct or incorrigible. This is a move in the right direction, but does not go far enough. No matter how much we change our concepts, as a result of learning new empirical facts or of philosophical arguments (like Gettier's), our concepts are still only contingently connected to metaphysical reality. In other words, primary/A meanings and secondary/C meanings are defined in terms such that they are only contingently identical; we can never assume them to be the same. No matter how much we refine our concepts, analyzing them alone does not give us *a priori* knowledge useful to metaphysics. Primary/A meanings can always fail to “cut the world at the joints”, such that they should not to be built into the study of the metaphysics of the actual world; if we stick to secondary/C meanings, if we stick to modalities which are indexed to or centered upon the actual world, we always then remain in touch with actual reality. The actual world, and not our concepts or vocabularies or accounts, should at all times be the touchstone of serious metaphysics.

We may, for example, “re-interpret” the folk concept of “free will” so that “it” is compatible with determinism, and thereby continue walking around saying that we believe in free will and determinism, but the metaphysical facts of the actual world have not changed a single iota: either agents have the power to act with metaphysical freedom, independent of the agent's and the world's past, or they lack such a power. We may change, as a result of Gettier, what we are willing to call “knowledge”, but whatever knowledge is, it has never changed. What we call things in the world is the business of language and communication, but it is not metaphysics. The analysis of our concepts, or our A/primary meanings, can at best yield to us analytically true *de dicto* results about how we conceive of things, it cannot yield to us true *de re* results about how the world is. And it is the latter which is the concern of serious metaphysics.

4. EPISTEMIC MODALITY AND IGNORANCE

In their joint paper (2001), Chalmers and Jackson shifted the formulation of their method by noting that the modality in question here, the modality built-into A/primary meaning, is epistemic in nature. We begin with what we perceive or what we already know and we ask what is logically consistent with this. Before the advent of chemistry, it was possible (as far as we

knew) for gold to have turned out to be something other than Au; it could have been some psychic material appearing to us as yellow, malleable, fusible, etc. We knew the stuff, whatever it was, appeared yellow, etc., and these facts alone are consistent with many possibilities. As soon as we know the actual truth about the stuff, we see that many of these so-called “possibilities” are actually impossible. Given the nature of the actual world, some “possibilities” are not possible.

The discussion of epistemic modality can begin where the technical innovations were developed to help us get clear on the distinction between centered and uncentered modalities. Stalnaker’s groundbreaking paper, “Assertion” (1978) can be our introduction to the “two-dimensional” modal framework. Stalnaker employs this framework specifically for helping us to understand what is going on during a *miscommunication* in which two of the three people involved have *false beliefs* about what is being said; they put a false interpretation on a sentence.⁹ We then can construct a two-dimensional model, built in terms of possible worlds, which allows us to understand what all the participants to the misunderstanding think is happening. We learn the important conclusion that semantic content is (in part) a function of context. What is relevant for our purposes is the fact that within this model, false belief is on the same footing as true belief; we consider counterfactual worlds at which beliefs that are false at the actual world are treated as “as actual” or “as true”. From a theoretical standpoint, within the two-dimensional model, true beliefs have no more warrant or authority than false beliefs. Given Stalnaker’s project of constructing models to help us understand semantic content, nothing could be more natural than this parity between true and false beliefs. On the other hand, for the purposes of serious metaphysics, and learning the metaphysical truth about the world, nothing could so easily lead us astray – the machinery involved was specifically designed to *not* distinguish between false beliefs and true beliefs. Using this machinery unnecessarily allows false beliefs into the dialectic of serious metaphysics.

Epistemic modality is the product of ignorance. The sense of “possibility” involved here is supposed to be the one reflected in the locution “for all I know”. As Chalmers says in the opening paragraph of a draft of a paper posted on the internet “The Nature of Epistemic Space”, “For all I know, it might be that there is life on Jupiter, and it might be that there is not. It might be that Australia will win the next Ashes series, and it might be that they will not”. “Possibility” in this sense is what is not ruled out by logic, given what we perceive or know. (Contrast this to the sense of “possibility” where it is what is not ruled out by logic, given what is actual.) Recall that the claim here is not that epistemic modality has no

proper philosophical uses, but is confined to the idea that it has no proper uses when investigating the metaphysics of the actual world. The reason for this is that when we take our perceptions or what we take as knowledge as the basis for our investigations we have no way left to distinguish veridical from illusory perceptions, and false justified beliefs from true justified beliefs. The mistakes in our thinking enter into our metaphysical considerations with the same justificatory status as what we understand correctly. “For all I know, it might be X and it might be $\sim X$ ”: both the truth and its contradiction have the same modal status here when, in fact, only one of them can actually be true.

When we are investigating the actual world itself, as opposed to investigating how we conceive of the actual world, this is a distinction which must be preserved, if only formally, to give us a chance of learning about our mistakes. If we allow epistemic modalities to inform our understanding of the metaphysics of the actual world, then we seal ourselves off from distinguishing when our intuitions about cases matter and when they do not. “For all I know, the members of the US Senate might be zombies and they might not be”. Well, given the actual world, we know there are no zombies (at least as well as we know anything at all), so the epistemic modality allows in the possibility that the Senate members are zombies when in fact this is impossible. The only avenues by which the possibility of zombies could be introduced into our consideration are either imagination or fancy (that is, ignoring the fact that we know that zombies are pure fiction) or our outright ignorance of the fact that zombies are actually impossible. Epistemic modality allows in the false along with the true, as long as the false does not contradict our experiences. Some of the possibilities which are true “for all we know” are actually false and some are impossible. Serious metaphysics should rule out these possibilities from consideration if it is at all possible and it is. A more rigorous sense of “possibility” than this should be employed, one which is centered upon the actual world instead of on our epistemic imagination.

None of this implies that when serious metaphysicians limit their attention to indexed modalities they thereby get results which are infallible. Mortal fallible humans cannot expect so much.¹⁰ For example, it used to be “known” *a priori* that space was Euclidean and now this appears to be empirically false. Still, by thinking that serious metaphysics employs an indexed sense of modality, we preserve a distinction between what possibly true “for all we know” and what is possibly true given what is actually true at the actual world. This helps us rule out what seems possible but is actually impossible and makes it more likely that we will discover the truth about the world.

5. A BRIEF AND CLOSING META-METAPHYSICAL PRESCRIPTION

For keeping metaphysics serious: stick to a discussion of those possibilities which are indexed to the actual world or to reality; stick to “secondary” or “C” meanings and extensions. Spurn those possibilities which only fail to be self-contradictory, those uncentered possibilities, and stick to actual possibilities. By doing so, we are more likely to learn what we set out to learn in doing metaphysics in the first place: we will learn about the nature of the single truly real and actual world. This is the only way to do metaphysics seriously.

NOTES

¹ “Let’s be realistic” is the title for a section of Hilary Putnam’s (1975) “The Meaning of ‘Meaning’”; “serious metaphysics” is a term of Frank Jackson’s *From Metaphysics to Ethics* (1998). I would like to thank the participants of the II Barcelona Workshop on Reference at which I first presented this work; conversations with Robert Stalnaker and David Chalmers were particularly helpful. I would also like to thank Crawford Elder, Michael P. Lynch, Thomas Bontly, and JC Beall for comments on an earlier draft of the paper.

² Quine’s “Two Dogmas of Empiricism” ([1951] 1953) is obviously in the backdrop of this whole discussion. While a full discussion would not be apt, it may be helpful for the reader to know that I take Quine’s arguments to show at most that (roughly) distinction between analytic and synthetic statements is vague. I do not conclude from this that there is no distinction at all to be had. See, Grice and Strawson (1956). Regarding causation, Crawford Elder pointed out to me after the paper was accepted that there is much current work being done on backwards causation which seems to contradict the spirit of what I say here. In response, if there is some unexpected reason to think that backwards causation is possible at the actual world, then theories of causation must account for it. If theorists are discussing backwards causation because it is logically possible, though actually impossible, then I think they are making a metaphysical mistake.

³ This may seem severe; a full picture will be given below to substantiate the claim. In the meantime, as a foreshadowing, consider how useful “primary” or “A” extensions should be for serious metaphysics when Jackson admits that in contrast to “C” extensions where “to know a term’s C-extension, we need to know something about the actual world”, we can “know the A-extension of ‘water’ at every world, for its A-extension does not depend on the nature of the actual world. Ignorance about the actual world does not matter for knowledge about the A-extensions of words” (1998, p.50).

⁴ Locke’s *Essay*, p. 439: “The measure and boundary of each Sort, or *Species*, whereby it is constituted that particular Sort, and distinguished from others, is what we call its Essence, which is nothing but that *abstract Idea to which the Name is annexed*: So that every thing contained in that *Idea*, is essential to that Sort. This, though it be all the Essence of natural Substances, that we know, or by which we distinguish them into Sorts; yet I call it by a peculiar name, the *nominal Essence*, to distinguish it from that real Constitution of Substances, upon which depends this *nominal Essence*, and all Properties of that Sort; which therefore, as has been said, may be called the *real Essence*: v.g. **the nominal Essence**

of *Gold*, is that complex *Idea* the word *Gold* stands for, let it be, for instance, a **Body yellow, of a certain weight, malleable, fusible, and fixed. But the real *Essence* is the constitution of the insensible parts of that Body, on which those Qualities, and all other Properties of *Gold* depend.** How far these two are different, though they are both called *Essence*, is obvious, at first sight, to discover". (Italics in original, bold face added here.)

⁵ Here I am using Chalmers' jargon as representative of the CJ view, but there is also something curiously backwards about Jackson's nomenclature as well. His "A intensions" and "C intensions" are analogous to Chalmers' primary meanings and secondary meanings respectively. But it is Jackson's "C intensions" which are specifically indexed to or centered on the Actual world, while it is in using his "A intensions" that we are supposed to be treating *non-Actual* or Counterfactual worlds as the actual world. As Jackson says, "To know a term's C-extension, we need to know something about the actual world By contrast, . . . [the] A-extension does not depend on the nature of the actual world" (1998, p. 50).

⁶ Putnam's "Meaning of 'Meaning'", p. 233 (note the modality is implicitly indexed to actuality): "Once we have discovered that water (in the actual world) is H_2O , *nothing counts as a possible world in which water isn't H_2O* . In particular, if a 'logically possible' statement is one that holds in some "logically possible world" *it isn't logically possible that water isn't H_2O* ". (Emphasis in original.)

⁷ Kripke's *Naming and Necessity*, note 13, p. 45: "Strictly speaking, Lewis's view is not a view of 'transworld identification'. Rather, he thinks that similarities across possible worlds determine a counterpart relation which need be neither symmetric nor transitive. The counterpart of something in another possible world is never identical with the thing itself. Thus if we say 'Humphrey might have won the election (if only he had done such-and-so), we are not talking about something that might have happened to *Humphrey* but to someone else, a "counterpart"'. **Probably, however, Humphrey could not care less whether someone else, no matter how much resembling him, would have been victorious in another possible world.** Thus, Lewis's view seems to me even more bizarre than the usual notions of transworld identification that it replaces". (Italics in original, bold face added here.)

⁸ I'd like to thank an anonymous referee for raising this defense of Chalmers.

⁹ Stalnaker writes "I said *You are a fool* to O'Leary. O'Leary is a fool, so what I said was true, although O'Leary does not think so. Now Daniels, who is no fool and who knows it, was standing nearby, and he thought I was talking to him. So both O'Leary and Daniels thought I said something false . . ." (emphasis in original, p. 317).

¹⁰ Indeed, it seems a likely truth of serious metaphysics that naturally evolved mortals are necessarily fallible, despite the epistemic possibility of the existence of evolved mortals who are perfect.

REFERENCES

- Block, Ned and Robert Stalnaker: 1999, 'Conceptual Analysis, Dualism, and the Explanatory Gap', *The Philosophical Review* **108**(1), 1–46.
 Chalmers, David: 1996, *The Conscious Mind*, Oxford University Press, New York.
 Chalmers, David: 2001, 'The Nature of Epistemic Space', at www.u.arizona.edu/~chalmers/papers/espace.html.

- Chalmers, David and Frank Jackson: 2001, 'Conceptual Analysis and Reductive Explanation', *The Philosophical Review* **110**(3), 315–360.
- Davies, Martin and Lloyd Humberstone: 1980, 'Two Notions of Necessity', *Philosophical Studies* **38**, 1–30.
- Gettier, Edmund: 1963, 'Is Justified True Belief Knowledge?', *Analysis* **23**, 121–123.
- Grice, H. Paul and P. F. Strawson: 1956, 'In Defense of a Dogma', *Philosophical Review* **65**(2), 141–158.
- Jackson, Frank: 1998, *From Metaphysics to Ethics*, Clarendon Press, Oxford.
- Kripke, Saul: *Naming and Necessity*, Harvard University Press, Cambridge.
- Lewis, David: 1968, 'Counterpart Theory and Quantified Modal Logic', *Journal of Philosophy* **65**, 113–126.
- Lewis, David: 1986, *On The Plurality of Worlds*, Basil Blackwell, Oxford.
- Locke, John: 1975 (1689): *An Essay Concerning Human Understanding*, Oxford University Press, Oxford.
- Putnam, Hilary: 1975, 'The Meaning of Meaning', in *Mind, Language, and Reality*, Cambridge University Press, Cambridge.
- Quine, W. V.: 1951, 'Two Dogmas of Empiricism', Reprinted in *From a Logical Point of View*, Harvard University Press, Cambridge (1953).
- Stalnaker, Robert. "Assertion" in P. Cole (Ed.), *Syntax and Semantics*, Academic Press, New York.

University of Connecticut
Department of Philosophy
344 Mansfield Road, U-54
U.S.A.
E-mail: phsb@uconn.edu