CAUSAL AND METAPHYSICAL NECESSITY

BY

SYDNEY SHOEMAKER

Abstract: Any property has two sorts of causal features: "forward-looking" ones, having to do with what its instantiation can contribute to causing, and "backward-looking" ones, having to do with how its instantiation can be caused. Such features of a property are essential to it, and properties sharing all of their causal features are identical. Causal necessity is thus a special case of metaphysical necessity. Appeals to imaginability have no more force against this view than they do against the Kripkean view that statements like "Gold is an element" are metaphysically necessary.

1.

There was a time, within the memory of some of us, when it was widely accepted that there are just two kinds of necessity. There was logical necessity, which was generally construed as including the necessity of analytic truths. This was assumed to be something to which we have an a priori access. And there was causal necessity, to which our only access is empirical. Since the dominant views about causality then current were Humean in inspiration, there was some question as to how so-called causal necessity, the distinctive status of causal laws and their consequences, deserved the name of "necessity" at all. But that was what it was frequently called.

Since then the boat has been radically rocked, first by Quine and then by Kripke. The intended effects of their attacks on the traditional view were of course very different. Quine's attack on the analytic-synthetic distinction sought to contract, if not to empty, the class of truths that are called necessary. Kripke, on the other hand, argued that the class of truths

Pacific Philosophical Quarterly 79 (1998) 59–77 0031–5621/98/0100–0000

deserving this label is much larger than had traditionally been supposed. And, in his most radical departure from the traditional view, he held that many of these truths have the epistemic status of being a posteriori. One important class of these truths included statements of identity, such as "Hesperus is Phosphorus" and "Water is H_2O". Another included statements about the essences of natural kinds, such as "Gold is an element" and "Tigers are mammals".

Such truths were characterized as being true in all possible worlds. That might suggest that they can be called logical truths. And that seems a radical departure from the traditional notion of logical truth. But so also, one might claim, was the earlier extension of the notion of logical truth to cover analytic truths such as "All bachelors are unmarried". If such truths are logical truths, they owe this status to the fact that their truth is guaranteed by certain paradigmatic logical truths – say that all unmarried men are unmarried – *together with* semantic facts, say that "bachelor" is synomymous with "unmarried male". What Kripke showed is that the class of semantic facts that can contribute to the bestowal of the status of necessary truth is much broader than the class of synonymies or analytic equivalences, supposing there are such; it includes such facts as that the term "gold" refers to a substance with a certain essential nature.

I favor the usage that restricts the term "logical truth" to what logicians would count as such, excluding both analytic truths like "Bachelors are unmarried" and Kripkean necessities like "Gold is an element". So I shall refer to the latter as metaphysical necessities, rather than as logical necessities. But this is not to say that metaphysical necessity is a weaker kind of necessity than logical necessity. A statement can be metaphysically necessary without being conceptually necessary, and without being logically necessary. This is true of "Gold is an element". But it is not plausible to put this by saying that there is a possible world in which gold is not an element, or by saying that the world might have been such that gold is not an element. We can compare this with the fact that "Bachelors are unmarried" is not, in the strict sense, a logical truth; we would not want to put this by saying that the world might have been such as to contain married bachelors.

So if Kripke is right, we have a set of truths, the metaphysically necessary ones, that are necessary in the strongest possible sense, and yet whose epistemological status is that of being a posteriori. Such truths are knowable, if at all, only empirically. And this epistemological status they of course share with statements of causal necessity. What, then, of the traditional view that causal necessity, if it deserves to be called necessity at all, is a weaker sort of necessity – that in possible worlds jargon, the causal laws hold in only a subclass of the metaphysically possible worlds? I think this view is widely held by philosophers who accept the Kripkean view that there are necessities a posteriori. But of course there is another

possibility. Maybe the view that causal necessity is a weaker sort of necessity, or necessity only in a stretched meaning of the term, is a holdover from the pre-Kripkean view, which took it that genuine necessities, or at least those of the strongest sort, are knowable a priori because a priori reducible to truths of logic. Maybe, instead, causal necessity is just a special case of metaphysical necessity, and is necessity in the strongest sense of the term.[1]

This is a view I defended in a paper published over fifteen years ago.[2] In that paper I put forward the view that properties are individuated by their causal features – by what contribution they make to the causal powers of the things that have them, and also by how their instantiation can be caused. Collectively, causal features of this sort constitute the essence of a property. So insofar as causal laws can be construed as describing the causal features of properties, they are necessary truths. One way to get the conclusion that laws are necessary is to combine my view of properties with the view of David Armstrong, Fred Dretske, and Michael Tooley, that laws are, or assert, relations between properties.

Views similar to mine have been defended by Chris Swoyer, and by Evan Fales.[3] But I think it is fair to say that it is definitely a minority view. The established view, even among those who have absorbed Kripke's lessons, is that causal laws are contingent, not just in the sense that their epistemic status is that of being a posteriori, but in the sense that there are genuinely possible situations in which they do not hold.

I shall not here repeat all of the arguments with which I supported the causal theory of properties and the necessitarian view about laws – although I will give a version of what I now take to be the central one. My main concern here is to clarify this view by exploring some of the sources of resistance to it.

2.

The source of resistance that most immediately leaps to mind lies in the fact that we can easily imagine what it would be like to experience a world in which the laws are different. We can imagine conducting crucial experiments and having them come out differently than they do here. And as Hume reminded us, we can imagine bread failing to nourish, water failing to suffocate, and so on.

But, of course, in the sense in which we can imagine these things we can imagine analyzing gold and finding that it is not an element, analyzing water and finding that it is not H_2O, dissecting tigers and finding that they are reptiles, and so on. Most philosophers are now persuaded, by Kripke's arguments, that such imaginings are no real threat to the claim that it is

necessary that gold is an element, that water is H_2O, and that tigers are mammals. Perhaps one wants to redescribe a bit what one actually succeeds in imagining; e.g., let it be analyzing what *looks like* gold, or what passes the layman's and jeweler's tests for being gold, and finding that it is not an element – or analyzing what passes the ordinary observational tests for being water and finding that it is not H_2O. Then one can grant that what one actually does imagine is possible, but say that it is wrong to describe the possibility as that of gold not being an element or water not being H_2O. Some would prefer to stick to the original description of what is imagined, but deny that imaginability establishes possibility. In any event, the same resources are available for someone who wants to maintain, in the face of all that we can imagine or seem to imagine, that causal laws are metaphysically necessary. If one's preferred strategy for dealing with the Kripkean examples is to challenge the claim that imaginability is proof or evidence of possibility, one can employ the same strategy here. If one's preferred strategy is to allow that imaginability is at least evidence of possibility, but to claim that what we can really imagine in these cases is not what we initially take it to be, that strategy too is available here.[4] Let the law be that strychnine in a certain dosage is fatal to human beings. We can grant it to be imaginable that ingesting vast amounts what passes certain tests for being strychnine should fail to be fatal to what passes certain tests for being a human being, but deny that this amounts to imagining a human being surviving the ingestion of that much strychnine.

But of course there was more to Kripke's argument than deploying such strategies to ward off challenges to his necessity claims that are based on imaginability or seeming imaginability. He had direct arguments for the necessity of identity propositions. He had a compelling case for the view that names and other singular terms are rigid designators. And, what is most pertinent here, he had a view of the semantics of natural kind terms that both implied that some statements expressed by the use of such terms are necessarily true and made it intelligible that such statements have the epistemological status of being a posteriori. It was part of that view that the semantic intentions underlying our use of natural kind terms are such that the underlying make-up of a natural kind, the properties responsible for the phenomenal features of the paradigm exemplars of the kind, is essential to it. So, for example, if the paradigm exemplars of gold are instances of an element having atomic number 79, it will be necessary that gold is such an element. This explains why something lacking that make-up can't be gold, whatever its phenomenal features. And it therefore explains why the imaginability of something having the phenomenal features characteristic of gold but lacking that physical make-up is no threat to the claim that it is necessary that gold has that make-up.

My claim will be that the causal theory of properties can provide a

parallel explanation of why the imaginability or seeming imaginability of worlds in which the causal laws are different is no threat to the claim that the causal laws that hold have the status of being metaphysically necessary. Corresponding to Kripke's claim that our semantic intuitions are such that the underlying natures of natural kinds are essential to them will be the claim that the causal features of properties are essential to them. Just as we can refer to a natural kind, and identify instances of it, without knowing what its essential underlying nature is, so we can refer to a property, and pick out instances of it, without knowing what most of its causal features are. Our ability to pick out instances of a property perceptually, or with the use of instruments, will of course depend on exercises of the causal powers which the property contributes to bestowing. For example, it may involve the power to produce experiences having a certain phenomenal character. But having the power to produce such experiences will not in general be sufficient for having the property in question; so the imaginability of something having such a power but lacking any property having certain other causal features is compatible with those other causal features being essential to the property in question.

3.

Before pursuing further the parallels between Kripke's claims and the claim that causal laws are metaphysically necessary, and before examining further the imaginability or seeming imaginability of the causal laws being other than they are, I must say more about what is, and what is not, asserted by the causal theory of properties I favor.

In expounding my theory in my paper "Causality and Properties" I made use of the notion of a "conditional power": something has a conditional power if it is such that it would have a certain power if it had such and such properties, where the possession of those properties is not itself sufficient to bestow that power. E.g., to use the example I used there, something that is knife-shaped has, among others, the conditional power of being able to cut butter if it is made of wood or steel. And I claimed there that properties can be identified with "clusters" of conditional powers, where the members of the cluster are causally unified in certain ways.

I now reject that formulation. One reason for rejecting it can be found in a postscript to that paper. I there consider a case offered by Richard Boyd, which purports to be a case in which two different properties bestow the same cluster of causal powers. We are to suppose that X is a compound of substances A and B and Y a compound of substances C and D, where A, B, C and D are all different substances, and that it is a consequence of the laws of nature that X and Y behave exactly alike in all possible circumstances – so being made of X and being made of Y bestow exactly

the same conditional powers. If, as seems plausible, X and Y are different substances, and therefore being made of X and being made of Y are different properties, we have a counterexample to the claim that properties are the clusters of conditional powers they bestow. This is not, however, a counterexample to the claim that it is essential to a property that it bestow a particular cluster of conditional powers. And if we include among the causal features of properties not only their "forward-looking" causal features, i.e., their being such as to contribute in certain ways to what their possessors cause, or what powers their possessors have, but also their "backward-looking" causal features, i.e., their being such that their instantiation can be caused in such and such ways, then Boyd's example is not a counterexample to the claim that properties are identical just in case they share all of the same causal features. In the example, the properties of being made of X and being made of Y share all of their forward-looking causal features, but not all of their backward-looking causal features.

Boyd's example raises another worry about the causal theory of properties, and I will return to it shortly. But even if his example could be shown to be impossible, and properties could be said to be necessarily coextensive with clusters of conditional powers, I would want to reject the formulation of the causal theory which says that a property *is* a cluster of conditional powers. That formulation has a reductionist flavor to it. And the reduction it seems to promise is a cheat. We must make use of the notion of a property in explaining the notion of a conditional power, so there is no question here of reducing properties to some more fundamental sort of entity.[5]

The formulation of the causal theory of properties I now favor is one that is in no way reductionist – and it is one that is compatible with the view that in Boyd's example being X and being Y are different properties. It says that the causal features of a property, both forward-looking and backward-looking, are essential to it. And it says that properties having the same causal features are identical.

You may well ask what exactly this theory is a theory *of*. Obviously it will not serve as a theory about the "abundant" properties recognized by David Lewis, where there is such a property corresponding to every set of possibilia, or every function from possible worlds to extensions. More plausibly, it is a theory about what Lewis calls "natural" properties, and what intuitively are "genuine", as opposed to "mere-Cambridge", properties. But for present purposes, we can take it as a theory about those properties that do contribute to determining the causal powers of things. Here I will be neutral as to whether there is a property common to all the things, future as well as past and present, that the predicate "grue" is true of. Even if there is, it is not this property that contributes to the causal powers of the things Goodman's predicate is true of. It is true that right now,

something will produce a certain visual appearance only if it is grue – so you can make it produce that appearance by making it grue. But that is because right now being grue goes with being green, and something's being green contributes to its producing that appearance. Or, if you are doubtful about the causal role of color properties, consider the predicate "negapos", which is true of something just in case it has negative charge and the time is before T (some time in the future), or it has positive charge and the time is T or after. Right now, things attract things with positive charge and repel things with negative charge just in case they are negapos. But, so I say, the property of them that bestows this power is the property of having negative charge, not the property, supposing there is one, of being negapos. Perhaps grue and negapos have causal features; but they have them only derivatively, in virtue of their relations to green and negative charge, respectively.

So the claim of the causal theory of properties is that the properties that have causal features *non*-derivatively have them essentially, and are individuated in terms of them.[6] I think this comes to much the same thing as saying that the properties that enter into causal laws have their causal features essentially, and are individuated in terms of them. I should observe that there is nothing to prevent a proponent of this view from saying that ordinary properties are not themselves dispositions, but are instead the "categorical bases", the having of which bestows on things the dispositions they have. Some writers, such as David Armstrong, appear to think that one can think of properties as categorical only if one takes it to be contingent that properties bestow the dispositions they do. But I see no reason to think this. There is, to be sure, a kind of dispositional property that is only contingently bestowed by categorical properties. To use Robert Boyle's famous example, it is only contingent that the intrinsic categorical properties of a key bestow the *extrinsic* power of opening a certain door, since their doing so depends on the door's lock having a certain structure and composition. But it is compatible with this that the *intrinsic* powers of the key, e.g., its capacity of opening locks of a certain structure and composition, are necessarily bestowed by its categorical properties.

In my original defense of the causal theory of properties I put considerable weight on epistemological arguments aimed at showing that there cannot be properties that make no contribution to causal powers, that there cannot be two or more properties that share all of their causal features, and that properties cannot change their causal features over time. These arguments have often been charged with being "verificationist". I think that charge is confused.[7] But in any case, my present formulation of the view makes it possible to dispense with most of them. If there are properties that make no contribution to causal features, the account does not apply to them. Since laws are supposed to hold timelessly, or omnitemporally,

we can be sure that the properties that figure in laws retain the same causal features over time. Where there is still room for the deployment of an epistemological argument is in attacking the idea that there can be properties that share all of their causal features. And what epistemological considerations show, in the first instance, is that if there are sets of properties whose members are identical with respect to their causal features, we necessarily lack the resources for referring to particular members of these sets. Supposing there are such properties, it cannot be these that we intend to be referring to when we use singular property-referring terms – all we can plausibly be said to intend to refer to are equivalence classes of such properties. So if there are such properties, they don't fall within the extension of *our* term "property". Which seems to imply that if there are such properties, they aren't properties; which seems to imply that there are no such properties.

At any rate, what divides proponents of the causal theory of properties from most other philosophers is not what they think about the *intra*-world identity conditions for properties. What divides them is what they think about the *inter*-world identity conditions. The controversial claim of the theory is that the causal features of properties are essential to them – are features they have in all possible worlds. That is what implies that causal laws are metaphysically necessary, and that causal necessity is metaphysical necessity.

4.

Now let me return, as promised, to Richard Boyd's example, involving the properties *being made of X* and *being made of Y* which are identical with respect to their forward-looking causal features. The first thing to be said about it is that it is not at all clear that the case is metaphysically possible. It is not clear that it is nomologically possible, and what is in question here is whether metaphysical possibility can outrun nomological possibility. The case is supposed to be one in which two substances, X and Y, have different compositions and yet are such that under no nomologically possible circumstances will they behave differently. This means that it is nomologically impossible for them to be decomposed into the different sets of elements of which they are composed. This is of course not true in ordinary cases in which different substances are phenomenally alike, e.g., in the case of the two sorts of jade, and it is at least questionable whether, given the physics and chemistry that actually obtain, it could be true of any pair of substances.

But let us waive this point, and assume that the example is possible. If it is, then there is a way in which the response I have already made to it, on behalf of the causal theory of properties, is inadequate. That response

involves saying that the causal features in terms of which a property is individuated include its "backward-looking" as well as its "forward-looking" causal features. Now it is true that it is compatible with the example that properties are identical just in case they share the same total set of forward-looking and backward-looking causal features. But how is it that we know that the property of being made of X and the property of being made of Y do not satisfy this condition? Is it because you get the first by combining the elementary substances A and B, and you get the second by combining the elementary substances C and D? That can't be it – there can be more than one way in which the instantiation of a single property can be caused, so for all that we have said so far it might be that there is a single property which can be gotten either by combining A and B or by combining C and D. Clearly, the intuition that being made of X and being made of Y are different properties stems from the intuition that X and Y are different substances. And the intuition that they are different substances stems from the presumed fact that they have different compositions. The presumption that they have different compositions of course comes from the way samples of the substances come into existence – namely, the fact that samples of the one are compounded out of the elements A and B, while samples of the other are compounded out of elements C and D. And the background assumptions operating here are not principles about the individuation of properties; they are assumptions about chemistry.

One might wonder if these assumptions would continue to be reasonable if cases of the sort Boyd imagines actually occurred; if some substances couldn't be decomposed into elements, it might be reasonable to think when they are formed out of elements those elements "fuse" in a way that destroys their identity – so that what is formed by combining A and B would not differ in composition from what is formed by combining C and D. But let's waive this worry, along with the earlier worry about whether the case is metaphysically possible at all, and assume that our A and B and our C and D are the elementary substances that make up, respectively, the compound substances X and Y.

What we now have to allow is that where properties are of the sort in question here, what we might call compositional properties, then there is a principle of difference that applies to them that does not have to do in any direct way with causality. The property of being made of x is different from the property of being made of y if x and y are different substances, and x and y are different substances if they differ in composition. This enables us to distinguish properties that are alike in their forward-looking causal features and that can be shown to differ in their backward-looking causal features only by the use of this principle. So it is not true without qualification that properties are "individuated by" their causal features.

As I have said, this in no way counts against the claim that properties are identical just in case they share the same total sets of causal features – it merely implies that non-causal considerations can come into deciding whether properties do in fact share the same total sets of causal features. And none of this is any threat to the claim of the causal theory of properties that especially concerns us here, namely that causal laws are metaphysically necessary. Arguably this follows from a part of the causal theory of properties that is in no way challenged by Boyd's example, namely that the forward-looking causal features of properties are essential to them. This implies that the laws about the effects of instantiating the various properties are metaphysically necessary. Moreover, if you hold fixed across worlds what properties there are, it is obvious that in fixing the forward-looking causal features of properties across worlds you at the same time fix their backward-looking causal features – just as in fixing what telephone calls are made by each of the members of the total population of telephone users one fixes what telephone calls are received by members of that population. So, again assuming that what properties there are is held fixed across worlds, if the forward-looking causal features of properties are essential to them, then all of the causal laws without exception, those about how property instantiations can be caused as well as those about what they can cause, will be necessarily true.

5.

But now I must address the question of why we should think that the causal features of properties are essential to them. I will approach this by considering first Kripke's case for the metaphysical necessity of statements about natural kinds. Since I am painting with a rather wide brush, I will here lump this together with Hilary Putnam's case for the same position, and draw on Putnam's examples as well as Kripke's.

Kripke asks us to consider a case in which, having discovered what we have discovered about the chemical make-up of gold, we discover, say on Mars, a substance that is phenomenally indistinguishable from gold, but which on chemical analysis turns out not to be an element at all, let alone one with atomic number 79. Elsewhere he asks us to consider creatures which look and act just like tigers but which biological analysis reveals to be reptiles, not mammals. And Putnam's Twin Earth contains a stuff which is just like water in its phenomenal properties, but which is XYZ rather than H_2O. We are expected to agree, and I do, that the Martian "gold" is not gold, that the reptilian "tigers" are not tigers, and that the Twin Earthian "water" is not water. And such examples seem to support the claims that the chemical make-up of gold and water, and the biological make-up of tigers, are essential to them.

But notice that what these examples have to do with in the first instance are the *intra*-world identity conditions for natural kinds. The world in which a gold-like substance on Mars turns out to be not an element is supposed to be one in which the actual samples of gold on earth are samples of an element with atomic number 79. And what Putnam imagines in imagining Twin Earth is not an alternative possible world but a remote planet in this world, or at any rate a world in which the watery stuff on Earth is H_2O. Now of course, it may well be nomologically impossible for there to be a compound that perfectly mimics the phenomenal properties of gold; what is called "fool's gold", namely iron pyrites, is very far from doing so. It may also be nomologically impossible for there to be reptiles that mimic the phenomenal properties of tigers, and non-H_2O that mimics the phenomenal properties of H_2O. I think that whether that is so is irrelevant to the philosophical point Kripke and Putnam are making. The question they address is whether, *supposing* that it is nomologically possible for there to be phenomenal duplicates of gold, tigers and water that differ from paradigm samples of these things in their underlying nature, these duplicates would count as instances of the kinds gold, tiger, and water. That question they answer in the negative – and expect us to agree with them in answering in the negative, on the basis of our sense of how natural kind terms of this sort are applied.

No one, to my knowledge, has replied: "Yes, it's true that given that the actual samples of gold or water have a certain underlying constitution, all samples of gold or water in the actual world, or in worlds in which the laws are the same as here, must have the same underlying constitution. But what of worlds in which the laws of nature are different? May not the underlying constitution of these kinds be different there?" Such a reply would seem strange, and I think we can see why. We go on the reasonable assumption that whatever restrictions there are on the variation of a kind in the actual world also restrict its variation across worlds. If we take it that there are worlds in which the laws are different, this will be seen as restricting the instantiation of these natural kinds to a subclass of worlds, namely those that are nomologically possible – those in which the laws are the same as here. Only in these can we sensibly speak of the same phenomenal features being produced by the same underlying constitution. But I think that there is a natural generalization of this line of thought that undermines the idea that this class of worlds – the nomologically possible ones – is a proper subclass of all the possible worlds.

It seems to me a general feature of our thought about possibility that how we think that something could have differed from how it in fact is closely related to how we think that the way something is at one time could differ from the way that same thing is at a different time. In possible worlds jargon, the ways one and the same thing of a given sort can differ across worlds correspond to the ways one and the same thing of that sort

can differ at different times in the same world.[9] Could I have been a plumber or an accountant instead of a philosopher? The answer seems to be yes – and this goes with the fact that we acknowledge the possibility of a scenario in which someone who was exactly as I was at some point in my life undergoes a series of changes resulting in his eventually being a plumber or an accountant. Could I have been a poached egg? *Pace* Lewis, the answer seems to be no – and this goes with the fact that our principles of trans-temporal identity rule out the possibility of a scenario in which something starts off as a human being of a certain description and ends up as a poached egg.

When we move from particulars to kinds, the relevant intra-world variation includes variation across space as well as variation across time. So, for example, the underlying constitution of gold cannot be different on Mars than it is here, just as it cannot be different tomorrow than it is today. Assuming that the constraints on intra-world variation are also constraints on inter-world variation, this gives us the Kripkean view that the underlying constitution of gold is essential to it, and belongs to it in all possible worlds.

But now consider the case of properties. Different things can be true of a property at different places, and at different times. E.g., the frequency of its instantiation can be different, and it can be coinstantiated with different properties, insofar as this is compatible with the laws. But it cannot be governed by different laws at different places or at different times. Applying again the principle that constraints on intra-world variation are also constraints on inter-world variation, we get the conclusion that the same property cannot be governed by different laws in different worlds. Since I take it that the causal features of properties are features they have in virtue of the laws that govern their instantiation, this is equivalent to my claim that the causal features of properties are essential to them. And it implies that causal necessity is a special case of metaphysical necessity.

6.

I want now to return to the point that we seem offhand to be able to imagine the causal features of properties being different, and, what goes with this, the laws of nature being different. I have already pointed out that a proponent of the causal theory of properties can respond to this point in ways that parallel Kripke's responses to the claim that, for example, we can imagine heat being something other than motion of molecules. But more can be said about it.

One natural thought is that for each of a wide range of properties there

is a perceptual "mode of presentation", and that when we take ourselves to be imagining something being true of certain of these properties what we are in the first instance imagining is that the thing in question is true of properties having that mode of presentation, and that that much, at least, is possible. If we could be sure that these modes of presentation necessarily always present the same properties, we could be sure that we are imagining what we seem to be imagining, and that it is indeed possible. What Kripke pointed out in his discussion of heat is that if we take as the mode of presentation certain sensations, there is no warrant for thinking that it is necessarily the case that these always present the same properties; in creatures with different sensory apparatus, the same sensations might be produced by something different than what produces them in us, namely motion of molecules, and in such a case the property they present would be something other than heat.

Pretending for the moment that perceptual modes of presentation are always sensations, there is one sort of property these are necessarily always modes of presentation of; sensations of sort S can always be taken to present the dispositional property of producing sensations of sort S under certain conditions. And it may be thought that we are at least in a position to establish via our imaginings the possibility that properties of this sort are governed by different laws than they are in the actual world. But this idea does not bear much examination. One can certainly imagine the regularities obtaining among one's sensations being very different than they in fact are. But obviously that is not to imagine any laws being different, for there are no laws that imply, by themselves, anything about what regularities obtain among sensations. Such laws as are relevant here will be ones about what sorts of sensations will be produced in a creature with certain sorts of sensory apparatus given certain sorts of stimulation from the environment. Imagining different patterns of instantiation among one's sensations will amount to imagining the breakdown of such laws only if the sensations can be taken as presenting the right sort of facts about one's own physical make-up and sensory apparatus, and about what sorts of stimuli are impacting on this physical make-up. And that is a very big "only if".

Now let me drop the fiction that sensations are the modes of presentation of perceptible properties. Indeed, I will drop talk of modes of presentation altogether. In its place I will adopt the following way of talking about the imaginings that are taken as establishing possibilities. What in the first instance one imagines, I will say, is a phenomenal situation. This phenomenal situation, in turn, is one whose existence is taken to indicate the existence of a state of affairs whose proper description is in "objective", non-phenomenal, terms. I use the notion of a "phenomenal situation" in a somewhat elastic way.[10] Sometimes it will involve the actual (as opposed to merely apparent) instantiation of what are naturally thought

of as directly observable properties – colors, shapes, etc. Sometimes it seems appropriate to construe the phenomenal situation as something more subjective, namely its appearing to one, perhaps over a significant interval of time, that such and such properties are being instantiated. In any case, the argument for possibility will go in two steps. What is in the first instance shown to be possible, by the imagining, is the phenomenal situation. The possibility of this, in turn, is taken to show the possibility of the more objective situation whose existence the phenomenal situation indicates. Perhaps there will be more than two steps – e.g., from the possibility of things appearing a certain way to the possibility of things *being* a certain way, in terms of directly observable properties, and from that to the possibility of some non-phenomenal state of affairs obtaining.

It is not only in the fantasies of philosophers that this sort of thinking goes on. It also goes on in scientific thinking. Entertaining a certain hypothesis leads one to consider what might be evidence for it. We can think of the possible evidence as a possible phenomenal situation. Having thought of, imagined, such a possible phenomenal situation, we think that the hypothesis may well be true. And to test whether it is true we set up an experiment having that phenomenal situation as a possible outcome. Or perhaps instead we are doing history of science, and considering some hypothesis that is incompatible with what we now know to be the case. We consider whether there are ways in which certain experiments in the past might have come out – phenomenal situations they might have produced – that would have established the truth of that hypothesis. We may well find that there are. So we conclude that something we now know to be nomologically impossible was possible.

Now I think it is apparent on slight reflection that in these cases the possibility of the hypotheses is epistemic. When we are considering the possible outcomes of an experiment we are about to conduct, and considering the different hypotheses we would take these outcomes as establishing or supporting, the hypotheses are possible relative to our current state of knowledge. When we consider a past situation, and say that a certain hypothesis, now known to be nomologically impossible, was then possible, we mean that it was possible relative to the then current state of knowledge. And I think that often when we say that a state of affairs is imaginable and so possible, what we really mean is that the thing is possible relative to some unspecified state of knowledge; we are considering some phenomenal situation we take to be possible, and saying that given some possible body of background knowledge, or justified background belief, knowledge of that phenomenal situation would make it reasonable to believe in the existence of that state of affairs. And I think that epistemic possibility, relative either to our actual state of knowledge or to some hypothesized state of knowledge, is all we need to account for cases in which we say that something is or was possible, where this is compatible

with the thing in question being nomologically impossible. What we do not need are metaphysically possible worlds which are not nomologically possible.

But what of the possibility of the phenomenal situations that figure in our imaginings? I think that there is a good case for saying that this is straightforward *nomological* possibility which we believe in on inductive grounds. We have abundant empirical evidence that when we can imagine some phenomenal situation, e.g., imagine things appearing certain ways, such a situation could actually exist – things really could appear that way to someone. In days of yore the evidence came from such things as the artistry of stage magicians and trompe l'oeil painters. Now this is supplemented by the productions of Hollywood special effects designers, holograms, "virtual reality" devices, and so forth. There is no need here, in the realm of phenomenal states of affairs, to suppose that metaphysical possibility outruns nomological possibility.

Of course, if it is on empirical grounds that we are entitled to think that the phenomenal situations we can imagine are nomologically possible, it should be imaginable that this should have turned out otherwise – that the empirical evidence should have supported the claim that some imaginable phenomenal states of affairs are nomologically impossible. And if that is so, it might be held, then it is imaginable, and so possible, that there should be phenomenal states of affairs that are possible, because imaginable, but not nomologically possible. And if it is possible that there should be possible states of affairs that are not nomologically possible, don't we after all need a notion of metaphysical possibility that outruns nomological possibility?

Let me observe first that there is a conception of imagination on which imagining a phenomenal state of affairs involves creating in one's mind an inner simulacrum of that state of affairs. To the extent that imagining is like this, imagining could be said to demonstrate the possibility of a phenomenal state of affairs by actualizing it. And in that case there is of course no question of the state of affairs not being nomologically possible.

But for most of us, most of the time, it is not like this. And I agree that it is imaginable, conceivable, that it should have turned out that what we can imagine in the way of phenomenal states of affairs outruns what is nomologically possible. E.g., experimental psychology might have discovered that there is a "speed limit" on certain sorts of phenomenal changes, which is such that these changes cannot occur as rapidly as we can imagine them occurring. But this, empirical psychology making such a discovery, is not itself a phenomenal state of affairs. And I say about its possibility what I said above about the possibility of scientific hypotheses that are not nomologically possible –there is no reason to think that this is anything other than epistemic possibility, relative to some envisaged body of background knowledge. So regarded, it gives no support to the idea that

imaginability gives us access to metaphysical possibility that outruns nomological possibility. It may further be observed that were this epistemic possibility to be realized then the imaginability of phenomenal states of affairs would establish nothing more than their epistemic possibility.

There is another sort of possibility that imaginability might be said to give us access to, namely conceptual possibility. By and large, we come to believe situations to be conceptually possible by reflecting on their descriptions and seeing no contradiction or incoherence. But the more theoretical the concepts involved in such a description, the less confident we can be that there is not a contradiction or incoherence that escapes our notice. And, conversely, the more the description is couched in phenomenal terms, the more confident we can be that there is not such a contradiction or incoherence. And what imagining a situation often amounts to is considering a description of it in phenomenal terms, or largely in phenomenal terms. Where the description is entirely in phenomenal terms, we will be entitled to regard the situation as not only conceptually possible but nomologically possible, given what I said earlier. When the description is not in phenomenal terms, we may nevertheless be justified in taking it to express a conceptual possibility. But as is well known, what is conceptually possible, like what is epistemically possible, may fail to be metaphysically possible.[11]

So there are three sorts of possibility we may reasonably come to believe in by reflecting on what we can imagine. There is the nomological possibility of phenomenal states of affairs. There is the epistemic possibility of non-phenomenal states of affairs. And there is the conceptual possibility of states of affairs of both kinds. But nowhere in all this, I submit, do we find any reason for supposing that there are states of affairs that are metaphysically possible but not nomologically possible.

7.

The causal theory of properties is so alien to what most philosophers believe that the burden of proof is commonly taken to fall on those who hold it. But I want to suggest, in conclusion, that once we accept with Kripke that there are metaphysically necessary truths whose epistemological status is that of being a posteriori, the burden of proof should be shifted to the other side. As is suggested by the way "necessity" is used in ordinary speech, and is defined in dictionaries, causal necessity is, pre-theoretically, the very paradigm of necessity.[12] The main theoretical obstacle to according it that status has been the deeply rooted conviction that necessary truths should be knowable a priori. Once that obstacle has been removed, anyone who holds that causal necessity is not really necessity, or has only

a second-class kind of necessity, owes us a reason for thinking this. I have not seen a reason that seems to me persuasive.[13]

The Sage School of Philosophy
Cornell University

NOTES

[1] Kripke explicitly recognized this possibility in *Naming and Necessity*: "Physical necessity, *might* turn out to be necessity in the highest degree. But that's a question which I don't wish to prejudge" (p. 99).

[2] See my l980.

[3] See Swoyer (1982) and Fales (1993).

[4] But if the position defended in this paper is correct, there are limitations on the applicability of this strategy – see note 11 below.

[5] I pointed this out in my 1980, but did not see then that it clashes with the reductionist flavor of some of my formulations.

[6] There is a case for saying that mental properties, and in general "higher-order" properties that are not themselves physical but which supervene on physical properties, have their causal features derivatively – that their causal efficacy derives from that of the physical properties in which they are realized. But there are ways and ways in which the causal features of a property can be "derivative"; it seems obvious that the causal efficacy of mental properties, and properties that figure in the "special sciences", is much more robust than the causal efficacy (if such it is) of the properties (if they are such) grue and negapos. The nature of the causal relevance and causal efficacy of higher-order properties raises large issues that I cannot go into here; but the causal theory of properties is meant to apply to at least some such properties.

[7] Those who make this charge, and they are legion, confuse verificationism (which involves commitment to a verificationist theory of meaning) with the use of a certain sort of epistemological argument, one that argues from the premise that we know facts of a certain sort (e.g., that two things share a property), together with the premise that we could not know them if such and such a view (e.g., that sameness of causal features is not sufficient for sameness of properties) were true, to the conclusion that that view is not true.

[8] I take the view stated here to be the standard view about laws. But it is sometimes suggested that it is conceivable that the laws might change over time; e.g., that certain physical constants involved in them might gradually change their values. Faced with an opponent who claims that that is a real possibility, I would fall back on my epistemological arguments. But I would also suggest that it may be that some of what we call laws are lawlike generalizations that hold in virtue of (a) "strict" laws and (b) boundary conditions that are highly stable but can (compatibly with the strict laws) undergo gradual change. If so, then it is not these laws, but only the strict laws, that I hold to be necessary and internally related to the properties involved in them. The suggestion that all laws hold in virtue of stable but mutable boundary conditions seems to me incoherent. If the stability of the boundary conditions were just a cosmic accident, it could not confer lawlikeness on generalizations. And to hold that the stability is the result of laws, but that these laws themselves hold in virtue of stable but mutable boundary conditions, generates an unacceptable regress.

[9] This principle needs to be qualified, as Randy Clarke pointed out to me. The property of being the child of someone who has visited Paris is not one that one can have and then lose; but it is one that one can have in the actual world and not have in some other possible

world. The principle is meant to apply only to non-historical properties; it says that where a non-historical property is not one that a given sort of thing can have and then lose, it is not one that a thing can have in the actual world and fail to have in some other possible world.

[10] Obviously my notion of a phenomenal situation is very similar to, and derives from, Kripke's notion of an "epistemic situation" – I have introduced my own terminology so as to be able to use it in a rather elastic way without getting entangled in exegetical issues.

[11] Some writers, taking their lead from Kripke's discussion of the seeming possibility of heat being other than molecular motion, have suggested that whenever we seem to be able to imagine something there is a metaphysically possible world that we are imagining, although it may be one we are misdescribing (as it is in Kripke's heat example). I do not disagree with this if it means that in those cases in which the imagining is in the first instance the imagining of what I have called a "phenomenal situation", the phenomenal situation, as opposed to the interpretation we put on it, is metaphysically possible. But some writers assume something stronger than this. They think that while we would be mistaken in taking Twin Earth imaginings to show that there can be water that is not H_2O, we are entitled to take them as showing that it is metaphysically possible for there to be a planet otherwise like Earth on which the "watery stuff" is not H_2O, where the watery stuff behaves, outside chemistry laboratories, just as water behaves here. If the view defended here is correct, this is a mistake. As noted earlier, it may well be that Putnam's XYZ, a substance that exactly mimics the behavior of H_2O except in certain laboratory conditions, is a nomological impossibility. And if it is a nomological impossibility, then according to me it is a metaphysical impossibility.

This may have the effect of undermining Kripke's argument against the idea that pain might be identical with some physical state, such as C-fiber stimulation. Granted that there is some sense in which one can imagine pain without C-fiber stimulation, or C-fiber stimulation without pain, this provides no guarantee that either of these situations is nomologically possible – and so, on the present view, no guarantee that either is metaphysically possible. If what one imagines here in the first instance is a phenomenal situation, no doubt that will be possible. It might be, e.g., feeling pain while having experiences "as of" a negative autocerebroscope reading (saying that there is no C-fiber stimulation going on in one), or having experiences "as of" a positive autocerebroscope reading while feeling no pain. But the metaphysical possibility of these situations is no proof of the metaphysical possibility of the situations whose possibility Kripke is out to establish.

[12] The first non-circular definition of "necessity" given in the *American Heritage Dictionary* (3rd edn, p. 1207) is "Something dictated by invariable physical laws". Much the same is true in *Shorter Oxford English Dictionary* (3rd edn, p.1391), where definitions in terms of "determined" and "constraining power" are followed by the definition "Constraint or compulsion having its basis in the natural constitution of things; *esp.* such constraint conceived as a law prevailing throughout the material universe and within the sphere of human action."

[13] This paper was written for the Kripke Conference held in San Marino in May l986. I am grateful for helpful comments from the participants of that conference, from audiences at the University of Iowa and Auburn University, and from members of John Heil's 1986 NEH Summer Seminar. I am also grateful to Eric Hiddleston for written comments on an earlier draft, and to Richard Boyd for discussion of these issues. It need hardly be said that I am greatly indebted to the work of Saul Kripke.

REFERENCES

Fales, E. (1993). "Are Causal Laws Contingent?" In Bacon, Campbell and Reinhardt (eds), *Ontology, Causality and Mind.* New York: Cambridge University Press.

Kripke, S. (1980). *Naming and Necessity*. Cambridge: Harvard University Press.
Shoemaker, S. (1980). "Causality and Properties", in Peter van Inwagen (ed.), *Time and Cause*. Dordrecht, Holland: Reidel, pp. 109–35.
Swoyer, C. (1982). "The Nature of Causal Laws", *Australasian Journal of Philosophy*, 60, 203–23.