

HOW THE MODALITIES COME TO INHERE IN THE WORLD

Wolfgang Spohn

Frege Lecture, Osnabrück, September 16, 2015

Dear Mr. President, dear Nancy and Simon, dear all of you who engage in the never-ending enterprise of philosophy!

And dear Hans, thank you so much for your many kind words! I can only humbly accept them. I am deeply moved, *deeply* moved and grateful for being awarded the Frege Prize. It is a true distinction! I am now the third to receive it; so the prize is still a young one. But it seems getting established. And I would be glad if it will turn out that I have contributed a little bit to its further establishment.

I would like to dedicate this prize to my two brothers. To my eldest brother Willfried, who died almost 4 years ago. He was an eminent scholar of historical sociology, the sociological and more theoretical counterpart of social history. As such he was always caught between two stools, sociology and history, but worked hard at, and succeeded in, establishing this field in Germany. I have always felt this to be the most relevant and fascinating branch of social sciences. And to my elder brother Herbert, who is presently about the most distinguished German mathematical physicist. Last year he received the Cantor Medal, the highest biannual award for a German mathematician. So, I confess that the Frege Prize is most welcome in view of the friendly competition within my family.

However, honors are not important, philosophy is. So, let's turn to it. Philosophy has partly taken the way of a normal science. In many respects this is a good and healthy development. Still, philosophy must never forget

about its grand questions; otherwise it would lose its heart. Let's plunge right into some of them.

The question I would like to discuss with you is, in a nutshell: where do the modalities come from? Modalities like metaphysical necessity, causation, nomicity, counterfactuality, dispositionality, probability and maybe more? We talk of them and make use of them everywhere, in everyday life and in all of the sciences. They seem very familiar. However, their nature is discussed only in philosophy; they are one of the few proprieties of philosophy; and they immediately go to the very heart of theoretical philosophy. For, they are deeply mysterious. They seem to belong to the world, but they are not before our eyes. They seem to refer to other possibilities, but other possibilities are nothing we can experience. So, how can they be *in* the world, as they seem to be?

I will suggest that we read those modalities into the world; they are projections or objectivizations of our epistemic modalities or attitudes, which we can understand very well. This is not a novel suggestion. It has forcefully entered the philosophical scene with David Hume and is since lingering in various forms of idealism, which sometimes pompously indulge in their revolutionary attitude and sometimes pretend to be quite innocent like Simon Blackburn's quasi-realism. Sure, projectivism is no less of a mystery; doubts are more than justified. What I would like to convey to you is that projectivism is more far-reaching than usually thought, that it thereby acquires greater coherence, and that it can indeed be carried out in constructive detail, which dissolves its metaphorical appearance and the appertaining doubts. Well, I cannot demonstrate the latter in this lecture, but the feasibility of constructive detail should at least become plausible.

My intentions may become more intelligible through the contrast with those of David Lewis. He had the same urge of not leaving the modalities unexplained. However, he favored a metaphysical explanation by designing his all-embracing so-called program of Humean supervenience, according to which all modal facts supervene on non-modal facts. I think this program

fails. Not only because it can ultimately not be carried to completion; one gets into fatal trouble with objective probabilities. But rather because it is misguided right from the start. In any case, what I would like to present to you is an even more comprehensive program, giving epistemology its dues.

Let me set out our issue in more vivid terms. Here we are, in that lecture hall or, if we slightly enlarge the environment, in our universe, the maximal space-time-like extension we belong to. In a sense, this is all there is. Well, in another sense, there are all the other possible worlds, as Lewis has claimed and I concur. If there are possible objects – and there certainly are: I could have had a daughter, though I don't –, then there also are possible objects with a maximal space-time-like extension, i.e., Lewisian possible worlds, or *universes*, as I shall call them. Lewis only erred in their status. Everybody followed him in calling them metaphysically possible worlds. This was not only to say that they are metaphysically possible objects; of course, they are. Rather, the label should suggest that they are well suited for accounting for metaphysical necessity and possibility. This, however, seems wrong to me. The universes are epistemic possibilities. The actual universe is our most comprehensive object of experience, though we don't know which one it is; it might be any of the possible universes, though we learn to exclude more and more of them. Thus they are our epistemic possibilities.

So, here is our universe. It is just one big, indeed maximal concrete object. However, the world is not just one big unstructured object. It is full of concrete objects. The objects have properties, they stand in relations, and thus they form states of affairs, some of which obtain. That is, the world is a world of facts. As Wittgenstein said: "Die Welt ist alles, was der Fall ist." Again, there is not only the actual world in this sense, other things might have had other properties, and other facts might have been the case. So, Wittgensteinian possible worlds are totalities of facts, or *totalities*, for short. And then they are not just totalities of ordinary facts; they seem to harbor modal facts as well, such as that I am necessarily human and pres-

ently produce sound waves, which reach your ear with nomic necessity and cause you to listen to me.

However, totalities belong to an ontological category entirely different from universes. Universes are concrete objects, whereas totalities are very complex states of affairs or collections thereof. The fundamental question is: How do we get from the one to the other? This is not a trivial matter at all; it rather is an enormous step that hides all our epistemological and metaphysical mysteries. It is in this step where we project our epistemic constitution in order to read the structure of a totality of facts into a universe. This is quite a Kantian idea. If you like, you may equate the actual universe with Kant's noumenal world and the actual totality with Kant's phenomenal world. This equation is apt in several ways. But it is also inapt in other ways, and so I shall not further dwell upon it.

I would like to divide this one enormous step into four still very large steps.

- Step 1: Why at all states of affairs? This, I will claim, is already a projection of our epistemic constitution.
- Step 2: How do objects enter states of affairs? I will explain that the notion of an object is already a modal notion, since it is tied up with the notion of an essential property. And I will explain that basically it is we who say what the essential properties of an object are and thus which objects enter into states of affairs.
- Step 3: Some states of affairs are facts, and others aren't. This is the basis of the correspondence notion of truth or its deflationary descendants. Sometimes, correspondistic truth is also called alethic modality. In my view it is a projection of the epistemologically more pertinent pragmatic notion of truth.
- Step 4: Here I will finally address, *pars pro toto*, one or maybe even *the* fundamental natural modality, namely causality, though only in a paradigmatic way indicating how it might fit into the overall projectivistic picture.

Let's start with the first step: Why at all states of affairs? Here, a simple state of affairs consists in a given object having a certain property or in a sequence of objects standing in a certain relation. And then there are complex states of affairs, which are algebraic combinations of simple ones. The question is: Why is it so natural, indeed entirely inescapable to conceive of the world in a Wittgensteinian way, i.e., as a collection, in fact a totality of states of affairs? How else could we conceive of it? You may indulge into an extravagant ontology, e.g., in an ontology of tropes. But it only assumes a novel kind of objects, which, though, form states of affairs again, by being similar and coinstantiated, etc. So, there does not seem to be any alternative.

You may say: well, this is the basic structure of the world; this is something we have to start from as a brute fact. Maybe; but this is only to concede that there is no further answer. Another popular response is that this basic structure of the world is a projection of the structure of our language. Sentences are basically structured in subject and predicate, or noun phrase and verbal phrase. We cannot help but thinking in terms of our language, and hence we cannot but endow the world with this structure. However, this leaves the basic structure of language as unexplained as the structure of the world before. At this point, Robert Brandom attempts to dig more deeply by arguing within his inferentialist semantics that there must be expressions with a symmetric inferential role and expressions with a non-symmetric role, i.e., subjects and predicates. I can't assess this argument here, but it will be clear that I am not satisfied and will take a very different route.

You may finally say: epistemology is the key that accounts for both, the structure of language and the structure of the world. I entirely agree. Yet, how does the key work? You may suggest that the contents of belief are propositions and that propositions have a conceptual structure and basically decompose into concept and individual concept. Then, however, your an-

swer is as short-circuited as before and tends to reversely project the structure of sentences into the structure of propositions.

However, the epistemological answer can indeed be carried one illuminating step further. I'll try to explain it as succinctly as possible. Another very common characterization of propositions is that they are simply truth conditions, i.e., sets of worlds. Such sets are unstructured and hence quite unlike states of affairs. If this is our epistemological starting point, it is hard to see how our explanatory strategy could carry us from such propositions to states of affairs. However, we well know that this characterization is insufficient. Egocentric attitudes or indexical beliefs cannot be accounted for in this way. Their representation requires a more liberal understanding of propositions, according to which they are sets of centered worlds consisting of a world and a center, which in turn consists of a subject (= I) and a time (= now). This is a step into the right direction.

It is little recognized, though, that this still won't do. Propositions are sets of epistemic possibilities, and an epistemic possibility is not just a centered world; we need to add a possibly infinite sequence of objects. The places of that sequence have various labels: mental objects, intentional objects, files, or addresses. They have quite an elusive existence; that's why they are often rejected. However, we can't avoid assuming them; that's why they are also often taken for granted. But what's the argument in their favor? The usual arguments I know are of a more or less explicit linguistic nature. Then, however, the present dialectic strategy does no better than assuming structured propositions right away, as we considered above. In my view, however, there is also a purely epistemological argument on behalf of amending epistemic possibilities by a sequence of objects. Very roughly, it runs as follows:

Objects are given to us under a mode of presentation according to Frege; we have only knowledge by description of them, says Russell. If that were so, if objects were epistemically represented by definite descriptions, we could stick to the first narrow sense of propositions as sets of centered

worlds. However, our concepts and the propositional structure generated by them may be more or less fine-grained. And the crucial point is that the descriptivistic picture is bound to presuppose a certain level of granularity. A larger set of concepts and thus a more fine-grained propositional structure may provide you with a unique definite description for the object you believe in, whereas a smaller set and a coarse-grained propositional structure doesn't; that is, your belief could not be represented in a coarse-grained structure. Thereby, the descriptivistic picture violates what I have called the *invariance principle*. I take this to be a basic principle of philosophical psychology and says that our static and dynamic descriptions and laws of propositional attitudes in general and epistemic attitudes in particular should be stated in a form that is invariant under coarse- and fine-graining of the propositional structure underlying those attitudes. To put it somewhat pompously, this means that the form of our mind should be independent of specific empirical contents.

Let me use a very simple example: You vaguely remember and thus believe that Kurt Gödel was a mathematician. Any conceptual structure able to represent this belief must at least contain the concept of a mathematician and an individual concept of Kurt Gödel. The invariance principle then entails that any conceptual structure containing these two concepts should be able to represent your epistemic dealings with the proposition that Kurt Gödel was a mathematician. The descriptivistic picture violates this, since it requires enriching the conceptual structure until it contains a definite description for Kurt Gödel you believe in. No such thing is needed according to the more complex epistemic possibilities proposed above. In their representation, you possess an individual concept of Kurt Gödel simply by having a mental file or address for him, one place in that possibly infinite sequence of objects amending an epistemic possibility. And thus the latter representation conforms to the invariance principle.

This seems to be a very abstract and hollow argument. Well, in the short time available I can't do better. However, one can dramatize the argument

by the familiar twin stories. Then one finds that your dynamics of belief concerning such hardly distinguishable twins cannot be adequately represented within the descriptivist picture using only centered worlds as epistemic possibilities. And one also finds that it can be adequately represented with the epistemic possibilities amended by mental objects. So, one might say that it is ultimately the dynamics of belief that enforces the amended representation. However, the core of those dramatized arguments is again the above-mentioned invariance principle, which has other fruitful applications besides the present one, in epistemology, decision theory, and the theory of causation.

I am aware that Kant seemed to aim at a similar conclusion when arguing for the concurrency of self-consciousness and object awareness in his refutation of idealism, which I do not understand well. Clearly, Brentano und Husserl had similar aims. I am not sure, though, whether they can be said to have proved the existence of intentional objects. And I am unclear about how their arguments are related to the present one.

Be this as it may, if the above argument is good, or can be made good, then epistemic possibilities are rather like models in the model-theoretic sense, and propositions, the objects of belief, are not assumed, but proved to be structurally like states of affairs. This finishes part of my first step. At least we have the right structure.

Still, propositions in this sense *are not* states of affairs. The latter are composed of objects, properties, and relations. However, concepts are not properties or relations, and intentional objects are not real objects; and not even possible objects; they are only epistemic placeholders for real or possible objects. So, we need to fill the places in the propositional structure with real or possible objects and with properties and relations in order to arrive at states of affairs. I will only discuss how objects, whether actual or possible, get into the picture. I think, though, that a similar account may be given of properties and relations. So, what are objects?

With this question we turn to the second step, the constitution of objects. I do not want to say that they are our epistemic projection; this would be the wrong metaphor. However, in my view it is we who cut up the one huge universe into pieces and read objects into it. How can I think so?

To begin with, we are very good in recognizing and identifying objects. Some animals are so, too, at least within their environment. However, this restriction certainly applies to us as well. Still, there is a crucial difference between animals and us. Like us, animals have the notion of qualitative identity, i.e., of falling under the same concept; this is all that it is needed for more or less reliable recognition. Like us, animals may also have mental files. However, only we humans have the notion of numerical identity, i.e., of identity proper and thus only we have the notion of an object proper. Animals don't. As Quine reminded us: no entity without identity. It's all very fine with discovering more and more evolutionary continuities. However, neither must one overlook the fundamental differences.

What, though, is numerical identity? Quine's saying really meant: no entity without identity conditions. Some philosophers seem to allow that objects have haecceities or thisnesses, non-properties, as it were, which secure the object's individuality. I confess, I do not understand haecceities: they are an utter mystery. What remains is to accept some version of Leibniz' principle of the identity of indiscernibles. Two numerically different objects must differ in at least one property they have.

Over which kind of properties do we quantify here? Properties deriving from identity, like being identical with me, must be excluded. Otherwise we would presuppose what we intend to explain. So, only proper properties, as I shall say, can be allowed. We well know that qualitative or intrinsic properties are not enough; in this sense there always are different, but indistinguishable objects. We must also allow proper relational properties. Those might also consist in spatiotemporal relations. And then, it seems, Leibniz' principle is easily satisfied.

However, and this is crucial, we must also exclude contingent properties. They are welcome for identification, but not for individuation or identity conditions. The problem is that there might be two different objects to which exactly the same properties and relations actually apply. There is me. Let's suppose we are roughly clear about my identity conditions. But there is also two-handed me, who is me with essentially two hands and who ceases to exist as soon as I lose a hand, whereas I cease to exist only when I die. Let's hope that I will never lose a hand. Thus, me and two-handed me actually occupy the same space-time-region and have actually all properties and relations in common. The only difference is that I have two hands contingently and two-handed me has two hands essentially.

I conclude that the identity conditions of an object are given by its essential, possibly relational proper properties. Two different objects must differ in one of those properties. I am well aware that I am treating here complex and difficult issues in an inadmissibly brief way. However, whatever the details, we must acknowledge the point that the identity of an object is bound up with its essence, the conjunction of its essential properties. And given this, we must ask: where does the distinction between essential and accidental properties come from?

Basically, I don't think we can find it in nature; it's not just there and waits for being discovered by us. We find a lot of properties instantiated in nature, but not whether they are essential or accidental. This distinction is fabricated by us. *We* tell which properties are essential for an object and thereby individuate or constitute that object in the first place. We can do it willfully, as I just demonstrated it by constituting two-handed me. But for the most part we simply connect up with ordinary language, which enshrines the ontological wisdom of our forefathers. Still, this means that the distinction depends on our linguistic conventions.

The point is obscured, but at the same time substantially deepened by what I call *Putnam's insight*. The insight is this: Prima facie it is clearly wrong what I just said. Sometimes we do discover the nature or essence of

objects and of kinds. It was a discovery that Venus is a planet and that water is H_2O . And didn't we thereby discover part of their essence? Yes, certainly; as Kripke has emphasized, there are metaphysical necessities a posteriori. However, Putnam described more clearly than others that it is still our convention to use our term "water" as a natural kind term, which heavily restricts the kind of essence we might find. We might as well have used it as we use "fire", where the superficial properties seem to be the essential ones; we call some phenomena fire even if there is no oxidization going on; fire is just a flamy appearance of light. In other words: our essentiality conventions need not completely fix the essence of objects or kinds, they can also leave room for empirical discovery. The latter clouds the conventional character, but does not make it vanish.

The fact that the constitution of objects is conventional is not to say that it is arbitrary. I cannot take any collection of properties and assume *the* object that has precisely those properties essentially; I cannot simply constitute *the* golden mountain as a concrete object by declaring these two properties to be its essential properties. You may think that the golden mountain makes sense as an incomplete object. However, I do not understand what an incomplete object is; possible objects are not incomplete. I confess I am not clear about the formal rules of object constitution. Neither am I aware that they have been deeply investigated, at least within the essentialistic framework presupposed by me. This would be something that, in my view, would truly deserve the Kantian label "logic of objects".

The formal rules are one thing. Our ontological policies are another. It would be most interesting to investigate why we constitute the objects and kinds as we do and not in any other way. By all means, we do not find out about this by doing physics, by exploring the nature of things ever more deeply. Rather we have to ponder about the rationality behind our essentiality conventions. Why do we constitute ordinary objects rather than their stages? Why rather enduring than perduring objects? Why natural kinds? I am not at all clear about that rationality. Again, I am not aware that it has

been well explored in terms of the rationality of ontological policy. Maybe those policies are somehow anchored in our explanatory policies. If plausible, this suggestion certainly needs a lot of spelling out.

I should finally remark that the conventionality of object constitution does in no way distract from the objectivity of the objects constituted. One must beware of modal confusions. Of course, the objects do not only exist when constituted; they also exist unconstituted; there are sundry of objects which are constitutable, as it were, but have never been constituted, have never been found worth attending by anyone. Surely, if I am pressed to give examples, I would first have to constitute them, as I did with two-handed me. However, two-handed me existed before I started talking of it. One must not make the mistake that our constitution of objects would be in any way essential to them. They would also exist if we never existed and never constituted them – unless, of course, we have produced them in the ordinary sense.

I mostly talked about objects and sometimes turned to natural kinds for exemplification. As I said, I assume that a similar account can be given for properties and relations. The basic notion for all this is that of a property being essential for an object or for another property. This metaphysical modality is involved already in ordinary objects, properties, relations, and states of affairs. And by fixing, or setting ranges for, what is essential for what we read metaphysical modality into the universe and thus generate all sorts of possible states of affairs. This concludes my second step and thus also my first step. The first step had provided only the structure of states of affairs,, and now the second step has filled the structure with content, with real objects, etc.

Let us turn to the third step. Some states of affairs obtain, are facts. More precisely, none of the states of affairs about non-actual, merely possible objects are facts, but half of the states of affairs about actual or real objects are facts. How, though, do some states of affairs come to be facts? This is the third step I want to consider.

Well, obviously this depends on how the world is, at least insofar as contingently existing objects are involved. If the world is conceived here as a Wittgensteinian world, as a totality, the issue is trivialized. A state of affairs is a fact relative to a totality if and only if it is contained in the totality, which is just an algebraically closed collection of states of affairs. And it is a fact simpliciter if it is a fact relative to the actual totality. If we thus have the facts, we can proceed to all the platitudes of the correspondence theory of truth and call beliefs, sentences, and utterances, true if they correspond to the facts. These are platitudes, basically because we use the very same words in the sentences and utterances as in describing the beliefs and the facts.

However, this is clearly not an account of how some states of affairs come to be facts. A totality is a totality of facts and thus presupposes such an account. So, if the facts depend on how the world is, we should rather conceive of the world as a universe, a Lewisian world. But what is a fact relative to a universe, what is truth relative to a concrete object such as a universe? This cannot be correspondence truth. Lewis did not seem to be aware of this problem, and neither did those working in his paradigm. However, there is a real problem. It is none less than: how is a universe transformed into a totality of facts? This is a grand issue, which you can enter only by taking David Lewis' comprehensive perspective, but which you cannot deal with within that perspective alone.

Since correspondence truth is of no help at this point, we have to refer, I think, to an alternative notion of truth, namely the pragmatic notion, which appeals to Peirce's idea of a limit of inquiry. I said that a universe is the most comprehensive object of experience for us. We know very well what it is to explore the world or parts of it; we do it everyday. So, we also have an idea of what it could mean to drive such an exploration to a limit where absolutely everything is explored. Of course, this limit is forbiddingly counterfactual or idealized. We cannot time-travel, we can actually look only at tiny regions of the universe, also because we live only for a minute

time. We do not have the cognitive capacities to record and process all the data. And we cannot observe the world and leave it unchanged at the same time. Still, I think that Peirce's limit of inquiry is not entirely ill conceived.

The pragmatic theory of truth then says that what we believe in the ideal limit of inquiry is true. Or as Putnam says: the ideal theory must be true. In the ideal limit there is no experience and no consideration left, there are no more reasons to be found that could falsify and change our judgments and beliefs. We are used to say that our beliefs aim at truth – quite a mysterious statement. In any case, there is an ineliminable gap between belief and truth; our a posteriori beliefs may always turn out to be false. Only in the ideal limit of inquiry this gap finally closes. It is in this way that we can understand the statement that our beliefs aim at truth.

You may tend to discount all this as nothing but nice metaphors. Of course, in the shortness of time, it can't be more. However, I am convinced that you can develop the pragmatic theory of truth so as to meet theoretical standards. The basic point is that you need a precise and general account of the change of our epistemic states, an account of the dynamic laws of belief. There are such accounts, not only mine. And once you have such an account, you can more precisely speculate about where our epistemic course leads to in the end, about the limit of inquiry. I am quite optimistic that we can thereby arrive at a proper independent and substantial theory of truth.

And I emphasize: an independent theory of truth. There are two theories of truth: an ontological one – this is the correspondence theory –, and an epistemic one – this is the pragmatic theory. A standard objection against the pragmatic theory and its appeal to the limit of inquiry or to the ideal theory is that the ideal theory is, of course, the true theory, true in an antecedently understood, presumably correspondistic sense. This would indeed defeat the entire project. However, it would mean to put the cart before the horse. No, we can partially characterize that ideal, by referring to epistemic rationality, which governs our epistemic states and their dynamics. And

reversely, epistemic rationality is substantiated through its entanglement with pragmatic truth. All this must and can be developed without any overt or hidden reference to the correspondence theory of truth.

A final important point: In the limit of inquiry not only our beliefs are guaranteed to be true; we also know then what the states of affairs and what the facts are. If we completely specify the identity conditions or the essence of an object, then we know right away which object we are thinking about. If, however, our essentiality conventions leave the identity conditions partly to nature, they can be known only a posteriori and hence for sure only in the limit of inquiry. For instance, if you say such a harmless thing as that having the parents I have belongs to my identity conditions, this makes for pretty demanding conditions reaching deep into earth history. In order to determine my identity, you have to go far back to the big bang and towards the limit of inquiry. Similar remarks apply to properties and relations. Do we know the essence of water? It seems so. Again, though, we can be sure only in the limit.

This explains how actual or only possible states of affairs, at least concerning contingent actual objects, are guaranteed to be determined only in the limit of inquiry. But then, of course, it is also determined which of the states of affairs are facts. And thus it is by carrying our epistemological enterprise concerning a given universe to the ideal limit that we transform this universe into a corresponding totality of facts. This is the upshot of my story so far.

Let me recapitulate: Why at all the structure of states of affairs? Because this reflects the structure of our epistemic states. And the latter structure must be as it is not in order to capture the structure of states of affairs, but because of our dynamic epistemic needs. Now the structure of states of affairs must be furnished with objects, properties and relations. They are not just there; it is we who must constitute them, i.e., define their essences or identity conditions. Why we constitute them as we do should somehow find a justification in our epistemic policies. Thus we get states of affairs. Some

of them are facts. But in order to determine the facts we have to carry our epistemic enterprise to completion and to appeal to the pragmatic theory of truth, which is profoundly entangled with our normative standards of epistemic rationality.

So, in some way or other, each step essentially depends on our epistemic constitution. This holds already for the non-modal part of the world in the sense of a totality of non-modal facts. However, this part is non-modal only because simple states of affairs are not *about* modality; they do not state something to be necessarily so. Still, as explained, already simple states of affairs presuppose a fully developed metaphysical modality, for which I have tried to specify the epistemic foundations.

Recall, moreover, my initial remark that I want to unfold an alternative program to David Lewis' program of Humean supervenience. Now you see that my program is even more comprehensive. The issues addressed so far are not dealt with or are taken for granted in Lewis' program. It is only now, in my continuation, that our programs have the same focus, although they proceed in a fundamentally different way.

For, I am convinced that the projectivistic strategy extends to modal facts in the ordinary sense, i.e., to the so-called natural modalities, which are also Lewis' subject matter. Let me paradigmatically demonstrate this, in my fourth and last step, with only one of the natural modalities, indeed the most important one, namely with causality. I confess that I am even more opinionated in this topic than in the former ones.

Concerning causality, our modern predicament starts with David Hume and his two theories of causation, none of which, by the way, is the counterfactual theory, as Lewis has suggested. The first is the well-known regularity of causation, where causal necessity boils down to constant conjunction. The second might be called Hume's associationist theory, according to which the effect is associated with the cause. Here, causal necessity is nothing but a customary transition or a habit of thought and hence an idea

of reflection, in Hume's terms. An idea of reflection – this cannot be over-emphasized! Kant's so-called Copernican turn originated from here.

Hume was ambiguous between the two theories and diminished their difference. Indeed, when he lets a critic object that the associationist theory is crazy, he swiftly resorts to the regularity theory. In fact, though, the difference could not be larger. The regularity theory is an objective ontological account of causation, and the associationist theory is a subjective epistemic account according to which causal relations depend on our associations. How the two theories relate is not easy to say. I think that Hume took the associationist theory to be conceptually basic and that he intended to explain the shape of our associations by the regularities, so that the associations agree with the regularities in the end. Thus he derived the regularity theory from the associationist theory.

The regularity theory has met so many objections that it appears untenable. And Hume's account of associations is definitely much too crude. Still I am convinced that Hume's scheme is basically right. Psychologists built a lot on Hume's theory of associations. As a philosopher I have a more rationalistic picture of them. I understand them as reasons; if I associate *B* with *A*, *A* is for me a theoretical or epistemic reason for *B*. And *A* is a reason for *B* if *A* speaks for *B*, or makes *B* more credible, that is, if *B* is more credible or less incredible given *A* than given non-*A*. This is my explication, the positive relevance notion of a reason. It is entirely subjective; what speaks for what is determined by my subjective epistemic state or, more specifically, by my conditional degrees of belief. I find the current debate about the nature of reasons quite confusing and quite confused, and I think that most of the other notions of a reason that are around can be captured on the basis of this subjective positive relevance notion and its various intersubjectivizations and objectivizations. That's another issue, though.

Now, causes are not just reasons, as my brief presentation of Hume might have suggested; whatever the confusing connection between reasons and causes, it is not that close. However, causes are a special kind of rea-

sons, indeed a special kind of conditional reasons. More precisely, according to my explication, A is a direct cause of B if and only if A is a reason for B given the entire history up to B without A . And causation, i.e., direct or indirect causation, is the transitive closure of direct causation. Since reasons are relative to an epistemic state, causation is so, too – just as in Hume’s associationist theory. A deep, but vague intuition concerning causation is that A is a cause of B just in case if I wiggle with A , i.e., intervene on A , B thereby wiggles as well. My explication agrees – it only interprets that wiggling epistemically. I have discussed the issue with Nancy Cartwright for more than 30 years, and I am well aware that she has forcefully attacked the common presupposition of the large debate to which I intend to contribute, namely that causation would be a single notion amenable to a uniform analysis such as mine or any other.

The main reason for my subjective turn is that in my view all objectivistic accounts of causation run into unsolvable difficulties. A nice case in point is symmetric causal overdetermination. In my hotel room the telephone and my mobile phone ring at the same time in order to wake me up at half past seven. Each of the phones would have sufficed to wake me up. So, this is a simple case of symmetric overdetermination. Such cases are ubiquitous. The basic difficulty of objectivistic theories with such cases is this: Somehow, a cause should make a difference to its effect; it should contribute to bringing about the effect. Each of the phones makes it true, by itself, that I wake up. However, in the presence of the other phone, each phone has no force; it cannot make my waking up doubly true, as it were. This doesn’t make sense; nothing can be truer than true. Hence none of the phones can be objectively positively relevant to my waking up in the presence of the other. Objectivistic theories are quite creative in order to overcome this difficulty, but nothing is convincing, I find. With subjectivistic theories the difficulty vanishes immediately. Of course, given that both phones ring you can expect me to wake up more strongly than given only one phone rings; you expect it doubly then, as it were. Epistemically,

hence, one phone can be positively relevant to my waking up in the presence of the other.

So, our subjectivistic starting point, the above explication, has much to recommend it. The crucial point, though, is that we are not stuck with it. Putnam wrote a famous paper: “Why reason can’t be naturalized.” He was right. In my somewhat different terms, reasons can’t be objectivized. That is, although unconditional beliefs can obviously be true or false, there is no general way to assign truth-values to conditional beliefs, and this applies all the more to the reason relation, which is based on a comparison of conditional degrees of belief.

However, Putnam extended his claim to causes, which he took to be of the same mold as reasons. Indeed, they are of the same mold according to my explication. Still he was wrong. As explained, causal judgments involve very special conditional beliefs: given the entire history up to t , what do I believe to happen at t ? These special conditional beliefs *can* be assigned truth-values after all, in a rigorous formal construction and under certain restrictions, which plausibly, but not necessarily obtain. I take those restrictions to be objectivization conditions of causal beliefs, which tend to be confused with defining characteristics of causation.

The details of this objectivization story are a bit complicated – and open to debate, of course. However, if the story is successful, it offers a constructive way of explaining how causal statements are true, not subjectively relative to an epistemic state, but objectively relative to the actual world. In this way – pretty much the way Hume had envisaged – there are causal states of affairs and indeed causal facts in the world. One may well call this a projection of the relevant conditional beliefs into causal facts. However, the metaphor should not conceal that it very much depends on the world whether and how the projection works. In any case, what counts is the above-mentioned objectivization theory behind the metaphor.

This was my fourth and last point. I feel I have to apologize for this overload of fundamental philosophy through which I rushed in a forbid-

dingly unphilosophical speed. However, the projectivistic strategy would be a mystery if it would apply only in an insulated way. It gains plausibility only by getting carried out widely and systematically. I did not apply it to moral matters, the other main concern of Simon Blackburn. But I did apply it broadly within theoretical philosophy: to truth, metaphysical modality, and causality. And I could have continued with nomicity, counterfactuality, dispositionality, and probability. My only hope is that by rushing through this wide range of topics the large explanatory potential of the projectivistic strategy and its constructive feasibility in detail has become apparent.

Did I try to convince you of queer idealistic conceptions? Do I suffer from a deficient sense of reality? I don't feel so. On the contrary, you may even think that I have been much too permissive concerning possible objects and also concerning actual, unconstituted, though constitutable objects (like two-handed me). There is only a problem with the projectivistic metaphor. It suggests that there is no projection without a projector. If there had never been a projector, there would have never been a projection and hence never been something projected. However, this causal connotation, though invited by the metaphor, is entirely inappropriate. Of course, the projected objects and facts, even causal facts, would exist also without anyone doing the projection. This is how our counterfactual talk works. The constitutable objects exist even unconstituted, even if nobody exists to constitute them. And the causal facts obtain, at least insofar they are accessible to the objectivization procedure indicated, even if there is no epistemic state around to be objectivized. Just as the ripe tomato is red, even if nothing like our color perception had ever evolved. It's only that there would be nobody around then to call it red and to conceive of it as red. This would be a new topic, though. In any case, rest assured that you have heard a real philosophical lecture.

Thank you very much for your long attention!