

Philosophical Papers

Volume II

DAVID LEWIS

NOTICE
This material may be
protected by copyright
law (Title 17 U.S. Code.)

New York Oxford
OXFORD UNIVERSITY PRESS
1986

PART 6: CAUSATION 157

21. Causation 159
 Postscripts to "Causation" 172
 22. Causal Explanation 214
 23. Events 241

PART 7: DEPENDENCE AND DECISION 271

24. Veridical Hallucination and Prosthetic Vision 273
 Postscript to "Veridical Hallucination and Prosthetic
 Vision" 287
 25. Are We Free To Break the Laws? 291
 26. Prisoners' Dilemma Is a Newcomb's Problem 299
 27. Causal Decision Theory 305
 Postscript to "Causal Decision Theory" 337
 28. Utilitarianism and Truthfulness 340
- Bibliography of the Writings of David Lewis 343
 Index 357

Introduction

Eleven of the papers in this volume were originally published from 1972 to 1981; misprints apart, they are reprinted in their original form. In some cases, where retractions or additions seemed urgently needed, I have appended postscripts. Two other papers appear here for the first time. The papers in this volume deal with topics concerning counterfactuals, causation, and related matters. Papers in ontology, philosophy of mind, and philosophy of language have appeared in Volume I. I have left out papers which are rejoinders, or which are of primarily technical interest, or which overlap too much with the papers I have included. Abstracts of the omitted papers may be found here, in the bibliography of my writings.

Many of the papers, here and in Volume I, seem to me in hindsight to fall into place within a prolonged campaign on behalf of the thesis I call "Humean supervenience." Explicit discussion of that thesis appears only in "A Subjectivist's Guide to Objective Chance"; but it motivates much of the book.

Humean supervenience is named in honor of the greater denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. (But it is no part of the thesis that these local matters are mental.) We have geometry: a system of external relations of spatio-

temporal distance between points. Maybe points of spacetime itself, maybe point-sized bits of matter or aether or fields, maybe both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated.¹ For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that.

First say it, then qualify it. I don't really mean to say that no two possible worlds whatsoever differ in any way without differing in their arrangements of qualities. For I concede that Humean supervenience is at best a contingent truth. Two worlds might indeed differ only in unHumean ways, if one or both of them is a world where Humean supervenience fails. Perhaps there might be extra, irreducible external relations, besides the spatiotemporal ones; there might be emergent natural properties of more-than-point-sized things; there might be things that endure identically through time or space, and trace out loci that cut across all lines of qualitative continuity. It is not, alas, unintelligible that there might be suchlike rubbish. Some worlds have it. And when they do, it can make differences between worlds even if they match perfectly in their arrangements of qualities.

But if there is suchlike rubbish, say I, then there would have to be extra natural properties or relations that are altogether alien to this world. Within the inner sphere of possibility, from which these alien intrusions are absent, there is indeed no difference of worlds without a difference in their arrangements of qualities.²

Is this materialism?—no and yes. I take it that materialism is metaphysics built to endorse the truth and descriptive completeness of physics more or less as we know it; and it just might be that Humean supervenience is true, but our best physics is dead wrong in its inventory of the qualities. Maybe, but I doubt it. Most likely, if Humean

¹ For ways to explain what makes a property natural and intrinsic, see my "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61 (1983): 343–77. However, I ought to add that besides the candidates considered there, class nominalism with primitive naturalness or a sparse theory of immanent universals, there is a third strong contender: a theory of tropes like that of Donald C. Williams, "On the Elements of Being," *Review of Metaphysics* 7 (1953): 3–18 and 171–92, but with the tropes cut to a minimum, so that the special status of natural properties is built into the ontology itself.

² On contingent supervenience theses, see the discussion of materialism in "New Work for a Theory of Universals." On inner and outer spheres of possibility, see Brian Skyrms, "Tractarian Nominalism," *Philosophical Studies* 40 (1981): 199–206; and D. M. Armstrong, "Metaphysics and Supervenience," *Critica* 14 (1982): 3–17.

supervenience is true at all, it is true in more or less the way that present physics would suggest.

I have conceded that Humean supervenience is a contingent, therefore an empirical, issue. Then why should I, as philosopher rather than physics fan, care about it? Isn't my professional business more with the whole expanse of logical space than with the question which of its districts happens to be ours?—Fair enough. Really, what I uphold is not so much the truth of Humean supervenience as the *tenability* of it. If physics itself were to teach me that it is false, I wouldn't grieve.

That might happen: maybe the lesson of Bell's theorem is exactly that there are *physical* entities which are unlocalized, and which might therefore make a difference between worlds—worlds in the inner sphere—that match perfectly in their arrangements of local qualities. Maybe so. I'm ready to believe it. But I am not ready to take lessons in ontology from quantum physics as it now is. First I must see how it looks when it is purified of instrumentalist frivolity, and dares to say something not just about pointer readings but about the constitution of the world; and when it is purified of doublethinking deviant logic; and—most of all—when it is purified of supernatural tales about the power of the observant mind to make things jump. If, after all that, it still teaches nonlocality, I shall submit willingly to the best of authority.

What I want to fight are *philosophical* arguments against Humean supervenience. When philosophers claim that one or another commonplace feature of the world cannot supervene on the arrangement of qualities, I make it my business to resist. Being a commonsensical fellow (except where unactualized possible worlds are concerned) I will seldom deny that the features in question exist. I grant their existence, and do my best to show how they can, after all, supervene on the arrangement of qualities. The plan of battle is as follows.

First, laws of nature. Few would deny that laws of nature, whatever else they may be, are at least exceptionless regularities. Not all regularities are laws, of course. But, following the lead of (a short temporal segment of) Ramsey, I suggest that the laws are the ones that buy into those systems of truths that achieve an unexcelled combination of simplicity and strength. That serves the Humean cause. For what it is to be simple and strong is safely noncontingent; and what regularities there are, or more generally what candidate systems of truths, seems to supervene safely on the arrangement of qualities. I stated such a theory of lawhood in my book *Counterfactuals*,³ and here I discuss it further

³ (Oxford: Blackwell, 1973).

in Postscript C to "A Subjectivist's Guide to Objective Chance."

I am prepared at this point to take the offensive against alleged unHumean lawmakers; I say there is no point believing in them, because they would be unfit for their work. Here I have in mind the theory that laws are made by a lawmaking second-order relation of universals, a theory most fully presented by D. M. Armstrong in *What is a Law of Nature?*⁴ Let N be the supposed lawmaker relation; the idea, in its simplest form, is that it is a contingent matter, and one not supervenient on the arrangement of qualities, which universals stand in the relation N ; but it is somehow necessary that if $N(F, G)$, then we have the regularity that all F 's are G 's. I ask: how can the alleged lawmaker impose a regularity? Why can't we have $N(F, G)$, and still have F 's that are not G 's? What prevents it? Don't try *defining* N in terms of there being a law and hence a regularity—we're trying to *explain* lawhood. And it's no good just giving the lawmaker a name that presupposes that somehow it does its stuff, as when Armstrong calls it "necessitation." If you find it hard to ask why there can't be F 's that are not G 's when F "necessitates" G , you should ask instead how any N can do what it must do to deserve that name.

Next, counterfactuals. I take them to be governed by similarity of worlds, according to the analysis given in "Counterfactuals and Comparative Possibility," in this volume. To the extent that this similarity consists of perfect match in matters of particular fact, it supervenes easily on the arrangement of qualities; and to the extent that it consists of (perfect or imperfect) conformity by one world to the laws of the other, it supervenes if the laws do. In "Counterfactual Dependence and Time's Arrow," I argue that one important sort of counterfactual, at least, will work properly if it is governed by just these respects of similarity.

Next, causation. In "Causation" and its postscripts, I defend an analysis of causation in terms of counterfactual dependence between events. The counterfactuals are discussed here in the two papers just mentioned; and since counterfactual dependence only seems causal when it is between events, my treatment of causation requires "Events" before it is done. Causation draws the arrow from past to future; that arrow exists only as an asymmetric pattern in the arrangement of qualities, so causal counterfactuals must somehow be sensitive

⁴ (Cambridge: Cambridge University Press, 1983). See also Fred I. Dretske, "Laws of Nature," *Philosophy of Science* 44 (1977): 248–68; and Michael Tooley, "The Nature of Laws," *Canadian Journal of Philosophy* 4 (1977): 667–98.

to the asymmetry. In "Counterfactual Dependence and Time's Arrow" I offer an account of that sensitivity. Given causation, or rather causal dependence, we can proceed to causal analyses of various things; for instance seeing, in "Veridical Hallucination and Prosthetic Vision," or what else you can do if you can freely raise your hand, in "Are We Free To Break the Laws?"

Next, persistence through time. I take the view that nothing endures identically through time. (Except universals, if such there be; their loci would coincide with relations of qualitative match, would indeed constitute these relations, so they would commit no violations of Humean supervenience.) Persisting particulars consist of temporal parts, united by various kinds of continuity. To the extent that the continuity is spatiotemporal and qualitative, of course it supervenes on the arrangement of qualities. But the continuity that often matters most is causal continuity: the thing stays more or less the same because of the way its later temporal parts depend causally for their existence and character on the ones just before. So the spatiotemporal boundaries of persisting things, for instance people, can supervene on the arrangement of qualities, provided that causation does. I discuss lines of causal continuity, not ruling out zigzag or broken lines, in "The Paradoxes of Time Travel." In "Survival and Identity," in Volume I of these *Papers*, I reply to some paradoxes brought against the idea that our survival is a matter of continuities that unite our temporal parts.⁵

Next, mind and language. Several papers in the previous volume concern the thesis that mental states, indexed with content when appropriate, are definable as the occupants of causal roles. Some of these states are people's beliefs, and some of their beliefs are their

⁵ It is at this point that Humean supervenience has come under direct attack. Saul Kripke, in "Identity through Time," given at the 1979 conference of the American Philosophical Association, Eastern Division, has argued that if a disk is made of homogeneous matter, then whether the disk is spinning or not is a feature of the world that does not supervene on the arrangement of qualities. We might have two worlds, just alike in their arrangements of qualities, one with a spinning disk and one with a stationary disk. (My "Humean supervenience" corresponds roughly to the "attenuated holographic hypothesis," which was one of Kripke's targets.) Whether the disk spins is, of course, definable in terms of the persistence of its parts through time; so in the first instance it is persistence that fails to supervene. But that might be because causation fails to supervene, and persistence requires causal continuity.

I reply by conceding, as I have, that Humean supervenience is contingent. The two worlds with their differing disks must (one or both) be worlds where there is something extra to make the difference. That does not show that any feature of *this* world fails to supervene on the arrangement of qualities. (Here I am indebted to Mark Johnston.)

expectations about other people. In *Convention*,⁶ and in “Languages and Language” and “Radical Interpretation” in Volume I, I suggested how semantic facts could obtain in virtue of the mutual expectations that prevail in a linguistic community.

And so it goes. There is room for endless argument over the details, but I remain confident that at every step mentioned the connection is something like what I have said—enough like it, anyway, to allow the cumulative Humean supervenience of one thing after another. At every step mentioned—but there is one that I passed over.

There is one big bad bug: chance. It is here, and here alone, that I fear defeat. But if I’m beaten here, then the entire campaign goes kaput. For chances enter at the very beginning. A law, I said with Ramsey, is a regularity that enters into the best systems. But what sort of systems? If there are chances—single-case objective probabilities, for instance, that a certain atom will decay this week—then some regularities have to do with chances, and the best true systems will be those that do best, *inter alia*, at systematizing the truth about chances. So bestness of true systems, and hence lawhood, and hence counterfactuals and causation and occupancy of causal roles and all the rest, will not supervene just on the actual arrangement of qualities, but on that plus all the chances there are, at various times, of that arrangement continuing in one way or another. Therefore the only hope for Humean supervenience is that the chances themselves might somehow supervene on the arrangement of qualities.

How could they? It is easy to go partway. The chances for alternative futures that obtain at a moment surely depend on just how things actually are at that moment. We might as well throw in the way things are at all previous times; that might help, and it’s no harm including too much. So far, so good. We have a conditional: if history is so-and-so then the chances are such-and-such. And the antecedent of that conditional—history up to the moment in question—surely does supervene on the arrangement of qualities.

But what is the status of the history-to-chance conditional itself? Is it necessary or contingent? If contingent, does it supervene or not on the arrangement of qualities?

If history-to-chance conditionals are necessary truths, no worries. Then the chances at any moment supervene on the arrangement of qualities, in fact on just the part of it up to that moment. Sometimes, we can see how the conditional might be necessary: suppose it says

that when we have prominent symmetry in the present set-up and its alternative futures, then those futures have equal chances. But sometimes not. How can an equality of chances based on symmetries, or any such necessary principle, give us the connections we need between, say, the exact height of a potential barrier and the exact chance of tunnelling through it? I hope there is a way, given the trouble I’m in if not; but I can’t see what it is.

If the conditionals are contingent, but themselves supervene on the arrangement of qualities, then again no worries. That would be so if they hold in virtue of relevant actual frequencies throughout the world, for instance. Or they could supervene in some fancier way, for instance by means of the “package deal” for chances and laws that I consider in Postscript C to “A Subjectivist’s Guide to Objective Chance.” Alas, I fear it cannot be so. The trouble is that whatever pattern it is in the arrangement of qualities that makes the conditionals true will itself be something that has some chance of coming about, and some chance of not coming about. What happens if there is some chance of getting a pattern that would undermine that very chance? The Principal Principle of “A Subjectivist’s Guide to Objective Chance” affords a way of turning this vague worry into a proper argument; hence the dismal ending to that paper.

Why not give in? I could admit that the history-to-chance conditionals, and so the chances themselves, are contingent and do not supervene on the arrangement of qualities. I could insist for consolation that at any rate all else supervenes on the arrangement of qualities and the chances together. Why not? I am not moved just by loyalty to my previous opinions. That answer works no better than the others. Here again the unHumean candidate for the job turns out to be unfit for its work. The distinctive thing about chances is their place in the ‘Principal Principle,’ which compellingly demands that we conform our credences about outcomes to our credences about their chances. Roughly, he who is certain the coin is fair must give equal credence to heads and tails; being less rough is the main business of “A Subjectivist’s Guide to Objective Chance.” I can see, dimly, how it might be rational to conform my credences about outcomes to my credences about history, symmetries, and frequencies. I haven’t the faintest notion how it might be rational to conform my credences about outcomes to my credences about some mysterious unHumean magnitude. Don’t try to take the mystery away by saying that this unHumean magnitude is none other than *chance*! I say that I haven’t the faintest notion how an unHumean magnitude can possibly do what it must do

⁶ (Cambridge, Massachusetts: Harvard University Press, 1968).

to deserve that name—namely, fit into the principle about rationality of credences—so don't just stipulate that it bears that name. Don't say: here's chance, now is it Humean or not? Ask: is there any way that any Humean magnitude could fill the chance-role? Is there any way that an unHumean magnitude could? What I fear is that the answer is "no" both times! Yet how can I reject the very idea of chance, when I know full well that each tritium atom has a certain chance of decaying at any moment?⁷

I thank all those who have helped me to think about these matters. Those who have helped me most are listed in the footnotes to the papers and the postscripts. Also I thank the University of California at Los Angeles; Princeton University; St. Catherine's College, Oxford; the American Council of Learned Societies; The University of Adelaide and the Australian-American Education Foundation; the National Science Foundation; Victoria University of Wellington and the New Zealand-United States Educational Foundation; Monash University; The Australian National University; La Trobe University; and all those universities that have given me opportunities to try these papers out on critical audiences.

For advice and assistance in planning these two volumes, and in seeing the project through difficult times, I am most grateful to Jim Anderson, Jonathan Bennett, Adam Hodgkin, Ruth Marcus, Tom Nagel, and Robert Stalnaker. I thank Nancy Etchemendy for the diagrams in Postscript E to "Causation."

D.L.

Princeton
October 1984

⁷ D. M. Armstrong has pointed out (in discussion) that matters are still worse if we grant that chances may take extreme values, one or zero exactly. Let H specify some course of history up to a certain moment and let F specify some course of history after that moment. Assume that H and F are contingent. (We need not assume that they are maximally specific.) Let T be a history-to-chance conditional which says that after history H , the chance of F would be exactly one. To grant that chances may take extreme values is to grant that some such H and T might both hold. Then is there any possibility that H and T might hold without F ? I say not. Any genuine possibility deserves at least some small share of credence, perhaps infinitesimal but not zero; but to give nonzero credence to this alleged possibility would violate the Principal Principle. So H and T strictly imply F . Now consider our three hypotheses about the status of history-to-chance conditionals.

1. Are they noncontingent? If so, T is necessary, since *ex hypothesi* it is at least possible. Then H by itself strictly implies F . How can that be? What prevents us having H without F , when they specify the character of wholly distinct parts of the world? This necessary connection between distinct existences is unintelligible.

2. Are they contingent, but supervenient on the arrangement of qualities? Then what would make T true is some pattern in the arrangement of qualities, and it is open to say that part of that pattern is simply that H does not hold or that F does. If so, we know how H and T can strictly imply F , so this second hypothesis gives no special problem about the case of extreme chances. But it still has its general problem: apart from the extreme case, how can a chancemaking pattern not give itself some chance of failing to obtain?

3. Are they contingent, and not supervenient upon the arrangement of qualities? Then if T is true, there is some unHumean feature of the world that makes it true. Call this unHumean chancemaker X . Now X and H strictly imply F . How can that be? How could X manage to impose this constraint on the arrangement of qualities? If we reject strict implication of F by H alone, as we should, then we grant that there are arrangements of qualities which make H hold without F . How does X prevent us from having any of these arrangements? Compare this unHumean chancemaker with Armstrong's unHumean lawmaker, denounced above. Armstrong has a fair *tu quoque* against anyone who accepts the one and balks at the other. For the two are alike in their supposed power to constrain the course of events, except that one imposes a connection in the single case and the other imposes a general regularity. (Indeed, the chancemaker might just be the lawmaker at work in one particular instance.) Either way, it's unintelligible how the unHumean constrainer can possibly do its stuff.

None of these three alternatives seems at all good. The escape routes from the trilemma—doubting that chances really can take the extreme values, doubting that every genuine possibility deserves some slight credence, or doubting the Principal Principle—seem just as bad. But so far as I can see, we must choose one evil or another.