Structural Realism and the Relationship between the Special Sciences and Physics

James Ladyman†

The primacy of physics generates a philosophical problem that the naturalist must solve in order to be entitled to an egalitarian acceptance of the ontological commitments he or she inherits from the special sciences and fundamental physics. The problem is the generalized causal exclusion argument. If there is no genuine causation in the domains of the special sciences but only in fundamental physics then there are grounds for doubting the existence of macroscopic objects and properties, or at least the concreteness of them. The aim of this paper is to show that the causal exclusion problem derives its force from a false dichotomy between Humeanism about causation and a notion of productive or generative causation based on a defunct model of the physical world.

1. Introduction. Scientific realists usually argue that we ought to be ontologically committed to the unobservable objects postulated by all of our best scientific theories, not only those postulated by our best fundamental physics. The standard form of argument for epistemic commitment to the existence of unobservable entities from tectonic plates to atoms is an inference to the best explanation, but what really does the explanatory work is not merely the existence of the objects in question but their causal powers to produce the observed phenomena. Since we only understand the entities as we do because of the way they are supposed to relate to our observations, belief in the unobservables entailed by our best scientific theories entails belief that those theories are at least approximately true, and that there really are the relations among the phenomena the theories attribute to the world. If those relations merely happened to be correlated with the unobservable objects in question, then positing the existence of the unobservables in question could not count as an explanation of the

†To contact the author, please write to: Department of Philosophy, University of Bristol, 9 Woodland Rd., Bristol BS8 1TB, UK.

Philosophy of Science, 75 (December 2008) pp. 744–755. 0031-8248/2008/7505-0021$10.00

phenomena and the relations among the phenomena.[1] This raises the two main issues with which this paper is concerned:

(1) the relationship between the ontology of the special sciences and the ontology of fundamental physics (recall Eddington's two tables);[2]
(2) the relationship between causation in the special sciences and causation in fundamental physics.

Issues (1) and (2) may have implications for each other. For example, if all macroscopic objects and properties were reducible to fundamental physical ones, then the causal powers of macroscopic objects would be reducible to those of microscopic objects. On the other hand, if there is no genuine causation in the domains of the special sciences but only in fundamental physics then there are grounds for doubting the existence of macroscopic objects and properties, or at least the concreteness of them.

Some forms of physicalism imply that everything that exists is physical; for example, Philip Pettit says the world "contains just what a true complete physics would say it contains" (1993, 213). Thus understood, physicalism is in tension with standard scientific realism. This kind of physicalism is also in tension with naturalism, since naturalists allow questions of ontology to be decided by science. Hence they ought to be realists about such entities as markets, fixed action patterns, mating displays, episodic memories, evolutionarily stable strategies, and phonemes, because successful explanation and prediction have been produced by special sciences that refer to such entities, and such success is sufficient for ontological commitment within science. Similarly, naturalism seems to demand that the causal claims of the special sciences be taken at face value, not least because, as mentioned above, it is the causal relations in which the entities posited by the special sciences feature that give us grounds for inferring their existence.

However, naturalists ought only to accept a form of physicalism that is motivated by reflection on the history of science and the nature and practice of contemporary science. Ladyman and Ross argue that this justifies nothing more than what they call the "primacy of physics constraint": "Special science hypotheses that conflict with fundamental physics, or such consensus as there is in fundamental physics, should be rejected

1. Entity realists contest this argument and play down the significance of theories of the unobservable, but they make even more of the causal powers of unobservable entities. So, for example, Nancy Cartwright (1983) argues for inference to the most probable cause as the form of argument that best establishes the existence of unobservables.

2. For the purposes of this paper, 'special sciences' should be understood as referring to all the sciences except fundamental physics (clearly some parts of physics are special sciences in this sense).

for that reason alone. Fundamental physical hypotheses are not symmetrically hostage to the conclusions of the special sciences" (2007, 44).

This leaves it open to the naturalist to believe both that the entities posited by the special sciences exist, and that the causal relations posited by them are genuine. However, the asymmetric methodological relationship between the special sciences and fundamental physics generates a philosophical problem that the naturalist must solve in order to be entitled an egalitarian acceptance of the ontological commitments he or she inherits from the special sciences and fundamental physics. The problem is that the causal exclusion argument, best known from the work of Jaegwon Kim in the philosophy of mind, can be generalized as follows: since fundamental physical causes are sufficient for all effects, and because there cannot be universal causal overdetermination, then there is no genuine causation described by the special sciences.

The aim of this paper is to show that the causal exclusion problem derives its force from a false dichotomy between Humeanism about causation and a notion of productive or generative causation based on a defunct model of the physical world.

2. Kim's Causal Exclusion Argument. Russell said that the idea of philosophy is to start with what everyone accepts and end up with what everyone will deny. This is not quite exemplified by Kim's causal exclusion argument, but he does manage to start with what a lot of philosophers accept, namely, nonreductive physicalism and the asymmetric supervenience of the mental on the physical, and arrives at the conclusion that there is no mental causation.

Kim's argument is effectively that the following propositions are inconsistent (modulo appropriate background assumptions):

(1) Effects are not generally overdetermined by causes.
(2) Mental states are not reducible to physical states.
(3) Mental states are realized by physical states and supervene on them.
(4) The physical world is causally closed (there is some set of physical causes that are sufficient for any physical event, or at least sufficient to fix its objective chance).
(5) Mental states can cause other mental states and physical states.

Kim argues that the premise that ought to be denied to restore consistency is (2). Once we accept that mental states are reducible to physical states, we can accept that mental causal relations are real just because they are reducible without residue to physical causal relations.

We ought to be immediately suspicious of the way the argument has been set up since Kim seems to have assumed the local supervenience of the mental on the physical. It is not clear that the physicalist ought to

accept anything more than global supervenience. Fred Keijzer and Maurice Schouten (2007) argue that the causal exclusion argument is undermined by what they call 'process externalism', the thesis that the subvenient base for mental states and processes at least sometimes includes parts of the body and the environment as well as the brain (the mind is 'extended'). Like Andy Clark and David Chalmers (1998) and Susan Hurley (1998), they claim that there is good empirical evidence for this view. It follows that sometimes the bearers of mental properties and the bearers of neural properties are different entities, and Kim does seem to think his argument relies upon the claim that mental states and neural states belong to the same entity (1998, 117). However, at best this argument only establishes that sometimes mental states are not epiphenomenal. Furthermore, disputing local supervenience does not get to the heart of the problem since we can reformulate the argument as follows. Suppose that M supervenes only on the whole physical state of the world at some time (or over some time interval), and likewise for M'; the former physical state of the world ought to be a complete cause for the latter physical state of the world, and since the latter realizes M', M is idle as a cause of M'. There are various other solutions that have been presented in the literature:

2.1. Accept Overdetermination. Obviously denying (1) offers an easy way out of the problem, and Kim's original attempt to block this option (1998, 45) has been criticized by others and modified by Kim (2005, 46–52). However, accepting the ubiquity of overdetermination without saying more might be to help oneself to a free lunch. The claim that effects do not have competing sets of sufficient causes seems innocuous enough until we ponder the exclusion argument, and some independent grounds for denying it seem necessary to avoid the charge of ad hocness. Nonetheless, on further analysis it may be that Ned Block (2003) and others are right to be sanguine about overdetermination. Perhaps when we analyze the cases that convince us that there cannot be overdetermination we will find that they are all cases where what is ruled out is overdetermination within a level, or within a special science. So, for example, it does seem that pervasive overdetermination within, for example, geology would be problematic and ruled out by geologists, but it is not so obvious that they would be bothered by the thought that geological events also have underlying physical causes.

Barry Loewer (2002) argues that Kim's objection to overdetermination depends on his commitment to a 'thick' metaphysical conception of generative or productive causation rather than a 'thin' notion of Humean causation as constant conjunction or counterfactual dependence. Kim seems to endorse this when he says that overdetermination makes little

sense when causation is understood in terms of "productive/generative mechanisms involving energy flow, momentum transfer, and the like, and not merely in terms of counterfactual dependencies" (2005, 47). Yet it is at least arguable that some kind of thick notion of causation seems to be required to defend scientific realism on the basis of inference to the best explanation. In any case, it would be interesting enough to learn that anti-Humeans about causation must deny mental causation or be reductionists about the mind. In Section 4 it is argued that the distinction between Humean causation and Kim's kind of generative or productive causation is a false dichotomy.

2.2. Program Explanation and Supervenient Causation. Frank Jackson and Philip Pettit (1990) propose the idea of program explanation, and attribute causal and explanatory relevance but not causal efficacy to the mental. Similarly, Kim himself (1984) used to advocate that mental causation was supervenient. Kim now rejects both these proposals on the grounds that they concede that mental causation insofar as it exists is a kind of epistemic construct that is explanatorily indispensable insofar as it gives us information about the genuine physical causation that underlies it, but which lacks the genuine causal "oomph" of the latter. As Pierre Jacob (2002, 651) says, Kim thinks that the problem of causal exclusion is about causation and not about explanation. Again, a Humean who disavows thick notions of causation is free to deny that there is a contrast class here, and hence to claim that once the explanatory and predictive 'causes' have been identified there is nothing more to know.

2.3. Cross-Classification. Terence Horgan (1997) points out that classifications of mental and physical causes in cases of nonreductive supervenience will differ with respect to their explanatory and predictive implications. This is easily seen in the case of a simple example. Suppose agent A believes that it is raining and this (among other things) causes them to take an umbrella out with them, while agent B does not believe it is raining and that this causes them to leave their umbrella at home. The subvenient base of A's and B's mentality might be very different at the neurological level. Hence, it is false to say of B that had they been in the same type of neurological state as A they would have taken an umbrella and vice versa. (Indeed, B could have in fact been in a similar neurological state to A.) Furthermore, it is not true to say of A that had they been in a different neurological state they would not have taken an umbrella since there may be other neurological states of them that would realize the belief that it is raining. It may be conceded to Kim that multiple realizability is limited and often overstated, but nonetheless it is clear that the explanatory and predictive power of the description of the agents at

the mental level is not available at the neurological level. When we consider explanation rather than causation, it is clear that explanations of why particular events happened are not available at the physical level.

To Kim (1998, 69) the assertion of an interesting form of cross-classification amounts to the denial of supervenience and implies a form of dualism. Perhaps here his argument does assume local supervenience in a way that matters. It may be that the state of all the mental properties is fixed by the complete physical state of the world, in keeping with global supervenience, even if more fine-grained specifications of the physical properties of brains do not suffice to fix the mental properties of people, in violation of local supervenience.

Kim says: "If we think of causation in terms of counterfactuals, we may assume that if P had not been there, the supervening M wouldn't have been there either, and that since M is what brought about P^*, P^* wouldn't be there either" (2005, 20). This seems false. Since M supervenes on P it is possible that had M not been realized by P it would have been realized by some other P' and that P^* would still have followed. Consider an analogy with thermodynamics. A gas expands and moves a piston. The increase in volume is caused by the pressure and temperature of the gas increasing. Now suppose that the physical microstate at the time when the gas begins to expand is P. It is not true that were it not in state P the piston would not move, because there are many possible physical microstates of the gas that are consistent with the same macroscopic properties of temperature and pressure.

Reflection on these three responses to causal exclusion arguments in the literature suggests that Kim's argument depends crucially on how we construe causation, and that the conflict the causal exclusion argument generates is with a particular metaphysical conception of cause rather than either Humean causes, or the operational idea of causation associated with everyday and scientific explanation and prediction. Of course, it is the latter that is actually tractable empirically. Before going on to see whether this offers the naturalist a way out, consider the problem generalized from the mental to the special sciences.

3. The Causal Exclusion Argument Generalized to Causes in the Special Sciences. Many critics (Fodor 1991; Baker 1993; Burge 1993; Bontly 2002) have argued that Kim's argument generalizes to one with the conclusion there is no macrocausation at all. This is then regarded, for example, by Block (2003) as showing that the original causal exclusion argument is unsound, if it is taken for granted that there is macrocausation and that interesting causal relations are described by the various special sciences. Burden-of-proof issues obviously arise here. Kim is clear that his point is not epistemological—we know that we and the objects of the special

sciences have causal powers—but metaphysical—if we are to make this belief compatible with physicalism we must embrace reductionism about the mental. It is open to him to argue that the generalized argument shows that we ought to be reductionists about everyday and higher-level scientific objects too (as he suggests in Kim 2005, 55). Note also that Trenton Merricks (2001) concludes from the falsity of such reductionism that the generalized causal exclusion argument shows that there aren't any macroscopic objects at all (except people) since he regards having causal powers as necessary for being taken as real.

However, Kim, perhaps recognizing that neither reductionism nor eliminativism about the kinds of the special sciences is plausible, offers another response to block the generalization argument. He appeals to the difference between orders of reality and levels of reality. Functionally individuated mental properties are higher-order properties that belong to the same entities as their first-order realizers, whereas physical properties of macroscopic objects are first-order properties of higher-level objects than microscopic objects. Hence, physicalism does not imply microphysicalism. Kim thinks that there are such things as 'micro-based macroproperties', the bearers of them "being completely decomposable into nonoverlapping proper parts" (1998, 84). There is no higher-order causation, but "macroproperties can, and in general do, have their own causal powers, powers that go beyond the causal powers of their microconstituents" (1998, 85).

More recently Kim says: "This baseball has causal powers that none of its proper parts, in particular none of its constituent microparticles, have, and in virtue of its mass and its hardness, the baseball can break a window when it strikes it with a certain velocity. The shattering of the glass was caused by the baseball and certainly not by the individual particles composing it. . . . the baseball = this composite structure of microparticles" (2005, 56).

If this notion of "going beyond" amounts to anything more than the sense in which epiphenomenal but irreducible properties go beyond their microconstituents, then it must mean that the instantiation of a first-order microbased macroscopic causal property does not supervene on the instantiation of the causal properties of the microentities. Indeed, Kim states this explicitly (2005, 57). However, this seems to introduce an inconsistency into his account, since if macroscopic objects with microbased macroproperties have nonsupervenient causal powers, then these threaten the causal closure of the physical level. After all, if a baseball breaks a window, then the positions of the microscopic particles of glass are caused to move and hence we seem to have downward causation, and a causally incomplete microlevel. Of course, Kim insists that higher-level causal relations between macroscopic objects are determined by and based on lower-level

causal powers plus the mereological relations between their bearers. Nonetheless, these mereological relations do not belong to the lower level and so closure is violated.

This aside, Kim claims that the difference between the mental causation case and the case of causation among macroscopic objects is that the bearers of these properties are not the same entities as their microscopic constituents. On the other hand, mental properties and neural properties are properties of the same things being "both had by human beings" (1998, 117). Mental properties are second-order properties functionally defined in terms of first-order physical properties. Kim seems to be a 'realizer' or 'filler' functionalist in *Mind in a Physical World*: "So *M* is now the property of having a property with such-and-such causal potentials, and it turns out that property *P* is exactly the property that fits the causal specification. And this grounds the identification of *M* with *P*. *M* is the property of having some property that meets specification *H*, and *P* is the property that meets *H*. So *M* is the property of having property *P*. But in general the property of having property *Q* = property *Q*. It follows then that *M* is *P*" (1988, 98–99). Similarly, in his "Making Sense of Emergence" (1999) he argues that there are only the first-order realizer/filler properties, and that what we call "second-order properties" are in fact concepts (17).

4. Dissolution and Diagnosis of the Causal Exclusion Problems. Kim says of a case of alleged mental causation: "there is only one causal relation here, namely a physical one, and more generally, causality is fundamentally a physical phenomenon" (2005, 55, note 22). He goes on to refer to Davidson's thesis that causation requires "strict laws" and that strict laws are found only in physics. Yet, as Clark Glymour pointed out in his review of Kim's 1998 book, it is arguable that there are no causal laws in physics and that the laws in physics are often not strict, and insofar as they are strict, laws in the special sciences are often as strict (Glymour 1999, 463). Whether or not there are or will be strict laws in fundamental physics, it is ironic that it is widely believed among philosophers who have looked carefully that there are no causes in fundamental physics. Bertrand Russell ([1913] 1917) famously diagnosed the commitment to a law of causation among philosophers as based on the projection into science of "anthropomorphic superstitions" deriving from the practical human predicament with respect to the asymmetry of past and future in our memories and capacities for control. More recently, Michael Redhead (1990) argues that laws of physics express functional relationships between physical quantities rather than describing causes, and that citing forces as causes is anthropocentric. Similarly, John Norton says, "mature sciences . . . are adequate to account for their realms without need of supplement by causal

notions and principles. The latter belong to earlier efforts to understand our natural world or to simplified reformulations of our mature theories, intended to trade precision for intelligibility" (2007, 12).

Of course, there is not unanimity about this among philosophers, and there are those, notably Cartwright (1989), who insist that fundamental physics ought to be construed as describing an ontology of causal powers rather than in Russelian terms. Nonetheless, this is a minority view, and without settling the matter it seems fair to say that it is at least an open question whether there is causation at the fundamental physical level. Hence, arguments to the effect that there are no causes in the special sciences that presuppose that all causation is at the fundamental physical level ought not to be taken too seriously.

Causal exclusion reasoning is further undermined by the fact that it is also an open question whether there is a fundamental level or an infinite chain of levels. Ned Block (2003) raises the problem that if causal powers of high-level entities drain down to the physical level, but there is no lowest physical level, then causal powers would drain away altogether. There may indeed be no fundamental level, and even if there is it may or may not be invested with causal powers by final physics. Furthermore, there may not be any levels at all contrary to the hierarchical ontological picture of the world that Kim and many others explicitly or implicitly presuppose.

In any case, Humeans of various stripes need not be troubled by causal exclusion. Some, like Loewer, construe causal relations in terms of counterfactuals, and they are then reduced to Humean facts about relations between possible worlds, whether concrete or abstract. Others will be pure regularity theorists who regard causal properties as at best secondary qualities. On neither account is there any reason to be worried about overdetermination. On the other hand, Kim's arguments appeal to those philosophers who believe that causation is an irreducibly intensional relation of necessitation (or some weaker modality), rather than being, or being reducible to, extensional relations between events or worlds. Unfortunately, Kim's view of productive or generative causation depends upon an unrealistic conception of the physical world: "By locating each and every physical item—object and event—in an all-encompassing coordinate system, the framework of physical space imposes a determinate relation on every pair of items in that domain" (2005, 85). That the structure of space-time need not admit of a single global coordinate system is one of the main innovations of the mathematics of manifolds that general relativity employs. Kim thinks about Humean causation in similar terms, regarding it as "an essentially spatial concept. Outside physical

space, Humean causation makes no sense" (2005, 86).[3] It is, of course, by no means clear that physical space will be regarded as ontologically fundamental by our best future physics, so if causation is fundamentally spatial then it is likely not be part of fundamental physics. The problem with Kim's metaphysical methodology is that it is based on intuitions about the world based on out-of-date science and folk physics: "The more we think about causation, the clearer becomes our realization that the possibility of causation between distinct objects depends on a shared space-like coordinate system in which these objects are located, a scheme that individuates objects by their "locations" in the scheme" (2005, 91).

Kim, like many metaphysicians, operates with what Ladyman and Ross (2007) call domesticated physics. Insofar as they discuss physical examples, they tend to describe the world in terms of what Lakoff and Johnson (1980) call the "containment metaphor." The world is supposed to be 'made of' myriad 'little things' in roughly the way that (some) walls are made of bricks. Unlike bricks in walls, however, the little things are in motion, and the paradigm of causation is the little things hitting each other. Hence, the causal structure of the world imagined by the domesticating metaphysicians is a network of 'microbangings'. The preoccupation with the search for 'genuine causal oomph' or 'biff' to settle the competition between different levels of reality derives from this conception of causation and microbanging. This is profoundly unscientific and does not fit with contemporary physics.

In Kim's 1998, despite its commitment to physicalism, there is no index entry for "physics" and no work of physics appears in the list of references. Kim's argument, however, depends on nontrivial assumptions about how the physical world is structured. One example is the above definition of a 'micro-based property' which involves the bearer "being completely decomposable into nonoverlapping proper parts" (1998, 84). Once again, the spatial conception of fundamental reality is illicitly presupposed. For Kim, microstructural constitution and levels are construed mereologically. On the other hand, composition in science, as opposed to in metaphysics and domesticated science, is dynamic and complex. Scientific explanations in terms of composition refer to distinctive features of specific domains rather than the metaphysician's illusory generic properties.

Note, however, that if there are no causal powers in fundamental physics, this does not imply that there are none in the special sciences. Harold Kincaid (see, e.g., his 2004) describes how the special sciences seek laws and theories of causal processes particular to their domains. Russell pre-

3. He claims that Hume included contiguity, but this was not a necessary condition, according to Hume, merely something that always seems to be present when there is a causal relationship.

supposes a conception of causation as "invariable uniformities of sequence" (1913, 178), and he is right that causation in this sense is not a feature of the scientific image of the world. However, the distinction between Humeanism and a thick notion of causation that is based on an outdated model of the physical world is not exhaustive. Ladyman and Ross (2007, Chapter 5) and Ross and Spurrett (2007) advocate a non-Humean naturalist view of causation that is compatible with both fundamental physics and the reality of causation in the domains of the special sciences.

REFERENCES

Baker, Lynne Rudder (1993), "Metaphysics and Mental Causation", in John Heil and Alfred Mele (eds.), *Mental Causation*. Oxford: Clarendon, 75–95.

Block, Ned (2003), "Do Causal Powers Drain Away?", *Philosophy and Phenomenological Research* 67: 133–150.

Bontly, Thomas (2002), "The Supervenience Argument Generalizes", *Philosophical Studies* 109: 75–96.

Burge, Tyler (1993), "Mind-Body Causation and Explanatory Practice", in John Heil and Alfred Mele (eds.), *Mental Causation*. Oxford: Clarendon, 97–120.

Cartwright, Nancy (1983), *How the Laws of Physics Lie*. Oxford: Clarendon.

——— (1989), *Nature's Capacities and Their Measurement*. Oxford: Clarendon.

Clark, Andy, and David Chalmers (1998), "The Extended Mind", *Analysis* 58: 7–19.

Fodor, Jerry (1991), "Making Mind Matter More", in *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press, 137–159.

Glymour, Clark (1999), "A Mind Is a Terrible Thing to Waste—Critical Notice: Jaegwon Kim, *Mind in a Physical World*", *Philosophy of Science* 66: 455–471.

Horgan, Terence (1997), "Kim on Mental Causation and Causal Exclusion", *Philosophical Perspectives* 11: 165–184.

Hurley, Susan (1998), *Consciousness in Action*. Cambridge, MA: Harvard University Press.

Jackson, Frank, and Philip Pettit (1990), "Program Explanation: A General Perspective", *Analysis* 50: 107–117.

Kim, Jaegwon (1984), "Epiphenomenal and Supervenient Causation", *Midwest Studies in Philosophy* 9: 257–236.

——— (1998), *Mind in a Physical World.* Cambridge, MA: MIT Press.

——— (1999), "Making Sense of Emergence", *Philosophical Studies* 95: 3–36.

——— (2005), *Physicalism or Something Near Enough.* Princeton, NJ: Princeton University Press.

Keijzer, Fred, and Maurice Schouten (2007), "Embedded Cognition and Mental Causation: Setting Empirical Bounds on Metaphysics", *Synthese* 158 (11): 107–125.

Kincaid, Harold (2004), "There Are Laws in the Social Sciences", in Chris Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*. Oxford: Blackwell, 168–185.

Ladyman, James, and Don Ross (2007), *Every Thing Must Go: Metaphysics Naturalised*. Oxford: Oxford University Press.

Lakoff, George, and Mark Johnson (1980), *Metaphors We Live By.* Chicago: University of Chicago Press.

Merricks, Trenton (2001), *Objects and Persons.* Oxford: Oxford University Press.

Norton, John (2007), "Causation as Folk Science", in Huw Price and Richard Corry (eds.), *Causation, Physics and the Constitution of Reality*. Oxford: Oxford University Press, 11–44.

Pettit, Philip (1993), "A Definition of Physicalism", *Analysis* 53: 213–223.

Redhead, Michael (1990), "Explanation", in Dudley Knowles (ed.), *Explanation and Its Limits*. Cambridge: Cambridge University Press, 135–154.

Ross, Don, and David Spurrett (2007), "Notions of Cause: Russell's Thesis Revisited", *British Journal for the Philosophy of Science* 58: 45–76.
Russell, Bertrand ([1913] 1917), "On the Notion of Cause", in Bertrand Russell, *Mysticism and Logic*. London: Unwin, 173–199.