Causal Mechanisms and the Philosophy of Causation
Ruth Groff

There is a way of thinking about philosophy, and about social science, such that it doesn't actually matter, for the purposes of doing social science, how philosophical questions are answered. In one sense, of course, it doesn't. If one is a realist, at least – and one ought to be – then one will think that the fact that one believes p, for example, implies nothing very much about the world other than that one does hold that belief. The world is as it is, one will say, regardless of which philosophical positions are correct. Even concept-dependent features of the world do not simply pop into existence via solipsistic command. But the "irrelevancy thesis," as I'll call it, is not an elliptical expression of realism. The idea that empirical inquiry is or should be unencumbered by the preoccupations of philosophers is a meta-theoretical principle, not a first-order ontological claim. I'm interested in the irrelevancy thesis as applied to the question of what causation is. Social scientists needn't care about such a question, say proponents of the thesis. If nothing else, competing answers can be true simultaneously. And besides, one can certainly put forward a causal explanation without holding any beliefs about what causation is. Lest I be accused of setting up straw men, let me say that these are positions that were defended with great conviction at a recent academic conference on the future of American sociology.

It is not very hard to show that the irrelevancy thesis is false. Manifestly, there are real differences between alternate accounts of causation (though the important ones are not between alternate iterations of Humeanism), and those theories that do genuinely differ from one another are patently irreconcilable. Moreover, insofar as a causal explanation is a causal explanation, it trades on a concept of causation, if only implicitly. And the approach that one takes to causation, even if one does so unreflectively, fixes what one may believe ***causes*** to be like, and by extension what sort of entity one will be in a position to think can even ***be*** a cause, as well as fixing what one will think is or is not happening when causation occurs. For these

reasons alone, apart from various others, it is simply not true that one can offer up a causal explanation without thereby having committed oneself to a view about what causation is, i.e., without implicitly having weighed in on a fundamental question of metaphysics. At this level, in fact, there is not much more to be said. One either abides by certain principles of discursive reasoning or one does not. But the irrelevancy thesis does not just lead its proponents to say foolish things at conferences that they may not really mean. Disregard for philosophy has also led to muddled thinking about so-called "causal mechanisms." And here, at least, there is something to be sorted out.

Towards that end, the discussion below is organized as follows: first, I establish a two-part typology that allows for a distinction to be made between passivist theories of causation and productive, anti-passivist theories of causation. In setting out each type of approach, I identify both how causation itself is defined from that perspective and what a cause must be, given the definition. I then turn to the notion of a causal mechanism, with the dual aim of illustrating existing philosophical confusion surrounding the term and of showing that philosophy can indeed help us to think more clearly about that to which the concept refers, if it does.

I. What is Causation?

Theories about what causation is fall into two broad categories. First, there are those views according to which causation is a kind of doing, a bringing about of change. Causes, from this perspective, really do produce effects -- literally. They do not simply precede them, or figure in theories about why they precede them. The idea that causation is productive assumes that activity, or dynamism, is an irreducible feature of the world. For if what looks like activity reduces ultimately to stasis, then causation could not really be about doing. It would only appear to be about doing, just as the figures in animated flip-books only appear to be animated. I will refer to the dynamic view of causation as a "powers-based" approach, and to the sense of

causation that it sustains as "productive" causation. Proponents of the powers-based approach take causes to be things (where "thing" is a count noun only, standing in for whatever kinds of phenomena might be thought to be efficacious) that are able to do x or y or z in virtue of being what they are, i.e., in virtue of their ability to "phi," as the analytic philosophers like to put the generic. Such properties are variously called "powers," "dispositions" and/or "dispositional properties" by those who understand causation to be productive. I should note that, increasingly, there are philosophers who use these same terms to describe elements of a passivist rather than dynamic ontology, and for whom, therefore, none of the terms is associated with genuine, irreducible activity. I will only use the terms in their normal, dynamic sense. Vis-à-vis causes who are persons, proponents of so-called "agent causation" often hold a powers-based view of causation, though not always. Similarly, with respect to inanimate causes, those who believe in "substance causation" may well see substances as productive causes. In the history of Western philosophy, Aristotle is the paradigmatic powers theorist, so that accounts of this type are widely considered to be neo-Aristotelian. However, Leibniz, Reid and perhaps Locke also fall into this category. Leading contemporary proponents of productive causation include: Brian Ellis[1]; Stephen Mumford and Rani Lill Anjum[2]; Nancy Cartwright[3]; Roy Bhaskar[4]; and Rom Harre and E. H. Madden.[5]

Second, there are those theories in which causation is ***not*** thought to be a matter of doing, producing or bringing about. I will refer to theories of this type as "passivist." Passivist theories themselves can be organized into three general categories. First, there are theories

---

[1] Brian Ellis, *Scientific Essentialism* (Cambridge: Cambridge University Press), 2001; *The Philosophy of Nature: A Guide to the New Essentialism* (Montreal and Kingston: McGill-Queen's University Press, 2002); *The Metaphysics of Scientific Realism* (Montreal and Kingston: McGill-Queen's University Press, 2009).
[2] Stephen Mumford and Rani Lill Anjum, *Getting Causes from Powers* (Oxford: Oxford University Press), 2011. See also, e.g., Stephen Mumford, *Laws in Nature* (New York and Oxfordshire: Routledge), 2004; Stephen Mumford and Rani Lill Anjum, "A Powerful Theory of Causation," in (ed., Anna Marmodoro) *The Metaphysics of Powers: Their Grounding and their Manifestations* (New York and Oxon: Routledge, 2010); Stephen Mumford and Rani Lill Anjum "Causal Dispositionalism," in (eds. Alexander Bird, Brian Ellis and Howard Sankey) *Properties, Powers and Structures: Issues in the Metaphysics of Realism* (New York and Oxon: Routledge, 2012).
[3] Nancy Cartwright, ( )
[4] Roy Bhaskar, *A Realist Theory of Science* (Routledge 4th ed)
[5] Rom Harre and E. H. Madden, *Causal Powers* (1975.)

according to which causation is a matter of rational or conceptual necessitation. From this perspective, causes and effects are phenomena that are linked necessarily, via some aspect or consideration of reason. Historically, Spinoza is an example of a rationalist who held this sort of view. As Michael Della Rocca puts it, "causation [for Spinoza] is nothing more than the relation whereby one thing explains another or makes it intelligible."[6] Though he is not a rationalist, Kant defends a variant of this type of passivism in the *Critique of Pure Reason*. As he has it there, causation is a transcendental principle of lawful order built into cognition itself. It is not that if we know the nature of a particular cause we can rationally deduce its effect, but rather that as rational beings we cannot help but experience the world as containing nomologically governed, non-contingent sequences of event.

A second, more prevalent type of passivism is associated above all with Hume. This approach also bottoms out ontologically in an assertion of order, but from this perspective the order in question is just a contingent a pattern or arrangement of given phenomena (for Hume, impressions), not anything to do with conceptual necessitation. Rejecting rationalism and Aristotelian dynamism alike, Hume himself equated causation, powers, natural necessity and all other ostensibly modal and/or dynamic phenomena with the feeling of expectation that accompanies impressions that, in the past, have been constantly conjoined with others. We project that feeling of anticipation onto the outside world, Hume said, which we then imagine to contain something called "causation" or "powers" or "natural necessity." Hume's approach has given rise to what Stathis Psillos terms RVC, the regularity view of causation.[7] Post-Humean proponents of RVC identify causation with the fact of regularity itself, not with a projected feeling of expectation, and they may well believe that it is something other than impressions that come in ordered sequences, but they retain the idea that a cause is not something that produces anything, or brings anything about. A cause is merely that which

---

[6] Michael Della Rocca, *Spinoza* (New York and Oxon: Routledge), 2008, p. 44.
[7] Stathis Psillos, *Causation & Explanation* (Montreal and Kingston: McGill-Queen's University Press, 2002; and "Regularity Theories," in in (eds., Helen Beebee, Christopher Hitchcock and Peter Menzies), *The Oxford Handbook of Causation* (Oxford: Oxford University Press, 2009), pp. 131-157.

always comes first, in an ordered pair. John Stuart Mill is an example of a post-Humean regularity theorist, as is Hempel. One might add Mackie to this list, at least insofar as he thinks of a cause as being a certain kind of sufficient antecedent condition, namely an INUS condition (an Insufficient but Non-redundant part of a condition that is itself Unnecessary but Sufficient).

An obvious problem for the regularity theorist is that the idea of constant conjunction does not do justice to our intuitive sense that there is an actual connection of some kind between causes and effects. Perhaps in response to this concern, some contemporary Humeans opt for counterfactual dependency theories over RVC. From a counterfactual dependency perspective, a cause is not just that which does comes first, but that which must come first, in the sense that had it ***not*** come first, the consequent would not have followed. Causes are thus antecedents that come first necessarily. In the *Enquiry*, Hume himself offered a re-formulation of his position in counterfactual terms: "Or in other words, where, if the first object had not been, the second never had existed." This said, Hume himself also claimed that "necessary" means only "what we have come to expect, given past experience," which is no doubt why he took counterfactual dependence and constant conjunction to be the same thing. But latter-day Humeans have undertaken to conceptualize necessity in a way that preserves the idea that there is a genuine difference of modal status between metaphysical necessity and metaphysical contingency. David Lewis, for example, defines necessity as that which is the case in all possible worlds, not just in this one. For this reason, it may indeed make sense to distinguish, at least provisionally, between Humeans who equate causation with the fact of regularity and those who equate it with counterfactual dependence. Certainly much is made of this distinction in the secondary literature. Ultimately, however, the appeal to possible worlds resolves back into regularity. Either the Lewisan move amounts to regularity being spread across a wider expanse (ontological space, if one is a realist about the possible worlds; logical space, if one is not), insofar as one has simply multiplied the number of worlds that display the same patterns of sequence that ours does; or it amounts to the deployment of a metaphor for the purpose of insisting that the constant conjunctions of this world really are constant. In either case, Humean

counterfactual dependency theories look very much like regularity theories, once we ask after the ground for the dependence. Crucially, proponents of both types of Humeanism deny the existence of irreducibly causal powers. They may or may not agree with Hume that the term has no meaningful content (except if we stipulate that it means regularity plus expectation), but they certainly agree that the world does not actually contain the phenomenon to which the term purportedly refers.

Finally, there is a third, somewhat curious form of passivism that is worth mentioning because of the attention that it has garnered in contemporary analytic circles, where it is often mistaken for a version of anti-passivism. I am thinking here of the view held by Alexander Bird, in which relations of genuine metaphysical necessity are thought to be built into the material world (via the essentialist nature of the identities of a restricted class of properties), but where these relations are not thought to be the stuff of causation, and where there is no other source or category of dynamism. Here too, what is salient for present purposes is the passivism. All passivists, regardless of whether or not they believe the world to be structured by, or to be necessarily experienced as being structured by, necessary connections -- all passivists agree that that there is no such thing as "dynamism" or "activity" that is not metaphorical.

The metaphysical divide between passivists and anti-passivists cannot be made to go away. Nor, despite analytic talk of "reducibly dispositional" properties, is it possible to have it both ways. The world is either fundamentally dynamic or it is not. As I have noted elsewhere, it may help to recall that this issue was at the center of metaphysical debate in the early modern period.[8] A key point of disagreement was whether or not God's on-going interference is needed in order for causes to bring about effects. Some parties to the debate thought no. God has endowed entities with their own causal powers. Others thought yes. Either (a) God's intervention is required because only God has causal powers; or (b) God's intervention is

[8] Ruth Groff, *Ontology Revisited: Metaphysics in Social and Political Philosophy*, Routledge, 2012, ch. 2. For helpful discussion see Walter Ott, *Causation and Laws of Nature in Early Modern Philosophy* (Oxford: Oxford University Press, 2009.

required because although God has endowed entities with their own causal powers, their powers alone are insufficient to produce change; or (c) God's intervention is required except for when human beings do things (because only God and human beings have causal powers). These are all viable positions. What would show up as plainly incoherent, by contrast, would be to say that while it is true that God created the world such that it contains its own causal powers, the truth-maker for the claim is a world does ***not***, in fact, contain its own causal powers. To put it in contemporary terms, there is no such real thing as dynamism that reduces analytically to something that is not-dynamism. The vote is up or down. The world contains displays of real powers-to-do, or it does not.

Given how difficult it is to avoid realist talk of doing – in discussions about causation and causes, to be sure, but also just in the routine use of the verb-form – it is understandable that those who are attracted to passivism might want to avoid coming to terms with the full brunt of what they are claiming about the world in disavowing the existence of real causal powers and/or irreducible dynamism. One form that this sort of equivocation can take (other than the idea of powers that reduce to non-powers) is that someone will defend a Humean account of causation, but unreflectively help themselves to a common-sense, anti-passivist conception of what it is to be a cause, which is to be that which brings something about. The other strategy, sometimes more self-conscious, is to reverse the order, i.e., to endorse an ostensibly dynamist or productive account of what it is to be a cause, but then fall back upon a Humean account of causation itself. The interventionist approach associated with James Woodward is an interesting example of this type of maneuver. Woodward defines a cause as that which has or could interfere in some way, thereby affecting a subsequent course of events. Here we are invited to picture the cause in productive terms, as something with powers. But it is not until we know what causation is, that we can know what it is to interfere. And on this score, Woodward tells us that interventionist theories are a species of counterfactual

dependence theory.[9] To cause is to necessarily come first. Thus the imagined counterfactual manipulation is itself causal not in virtue of involving real causal powers, but rather in virtue of the same fact of order that is thought to render causal the framing counterfactual relation in which it stands.

Bearing in mind, then, the distinctions between passivist and anti-passivist accounts of causation, and between the question of what causation is and the question of what it is to be a cause, let me turn now to the literature on causal mechanisms.

II. Causes, Causation and Causal Mechanisms

Broadly and provisionally, I will assume that by causal mechanism those who use the term have in mind something like "The missing steps between x and y, when x is the cause of y" or otherwise "The means by which x causes y." For the moment, the difference between these two formulations doesn't matter. The key idea is that of a mediating explanatory element, formerly overlooked, which properly connects cause-x to effect-y. Some authors describe this bridge in terms of the so-called black box of causation: a causal mechanism is what we see when it turns out that we can look inside. Other times is it conceived in terms of the tricky little conjunct "whereby." Those who invoke causal mechanisms see their approach as differing fundamentally from that of those who take explanation to be subsumption under a law. Whatever it is that we learn from "Y occurred because 'When x, y' is a law," it obviously does not tell us ***how*** x caused y. In addition, use of the term "mechanism" suggests that, *contra* proponents of a covering law approach, those who invoke causal mechanisms do not endorse the passivism of constant conjunction. The choice of metaphor signals that something gear- or spring- or trigger-like is imagined to be operating, when causation occurs.

[9] James F. Woodward, "Agency and Interventionist Theories," in *Op. Cit.*, (eds., Beebee, Hitchcock and Menzies), *The Oxford Handbook of Causation*, pp. 234-262.

What are we to make of this alternative to a Hempelian-derived approach? Is it really an alternative? Is it cogent? If we disregard the advice to ignore philosophy, and opt instead to assess the causal mechanisms literature with an eye to its implicit metaphysics, what we find are (i) accounts of what it is to be a causal mechanism that don't add up either because their proponents are anti-realists about the posited mechanisms or because they are anti-realists about causation itself, or both; (ii) accounts that don't add up because they involve a conflation of causation and cause (either what it is to be a cause or what causes what in a given case or type of case), and perhaps other confusions as well; and (iii) accounts that are coherent but that nevertheless raise important questions about the very concept of a causal mechanism. Surely, it seems to me, these findings will be of concern to sociologists, and not just to philosophers.

(i) Anti-Realism

There are two different ways to be an anti-realist about causal mechanisms. The first is to be an anti-realist about the posited entity or process that is the mechanism. Whether it be construed as a set of variables, a sequence of events or states of affairs, a system or some other complex whole -- regardless of the conception, the ontological status of a causal mechanism according to someone who is an anti-realist at this level of abstraction is that the mechanism itself is "a bit of theory," a "piece of scientific reasoning," as Stinchcombe puts it.[10] Notice that there are nested layers of metaphor at play here. Patently, any "mechanism" that is just a heuristic is a mechanism only metaphorically. This will be so regardless of how it is conceptualized. But for many anti-realists, even the mechanism-that-doesn't-really-exist is just a metaphor for something else that doesn't really exist. It's not even an ***imaginary*** mechanism; rather, it's an imagined set of intervening variables, metaphorically ***cast*** as an imaginary mechanism. Indeed, it's worth distinguishing between such a view and the view that while mechanisms are nothing more than variables, the variables themselves refer to something real.

[10] Arthur L. Stinchcombe, "The Conditions of Fruitfulness of Theorizing About Mechanisms in Social Science," *Philosophy of the Social Sciences*, Volume 21, Number 3 (September 1991): 367-388, p, 367.

Often the latter position is also considered to be anti-realist by those who believe mechanisms to be real entities or processes. For present purposes, I am going to set aside all anti-realist theories in which it is the posited mediating phenomenon whose reality is denied. If causal mechanisms are not real features of the world, then confusions surrounding their ontology are neither here nor there. The only point to be made, one that I have made in full elsewhere, is that a mechanism that doesn't actually exist also hasn't actually been the cause of anything.[11]

The second way to be an anti-realist about causal mechanisms is to be anti-realist about causation itself. Here the claim is not that the posited apparatus doesn't really ***exist***, but rather that causal mechanisms, in keeping with a passivist metaphysics, can't actually ***do*** anything. This is the version of anti-realism worth discussing further, and it is a position in relation to which philosophical clarity is crucial. I say this for two reasons. First, it is increasingly common for those who take causal mechanisms to be something other than intermediary variables to be attuned to the question of the ontological status of the invoked phenomenon, and to claim the title "realist" for their approach in virtue of thinking (a) that mechanisms are real (as opposed to being a mere heuristic or posit) and (b) that they are entities, steps or processes, rather than variables. But matters are not so simple. To be sure, the "realist" about causal mechanisms who is a passivist when it comes to causation holds a view that differs from that held by the instrumentalist about mechanisms. She thinks that mechanisms are real, while the instrumentalist thinks that they are not. But the fact that she doesn't believe causation ***itself*** to be real, i.e., the fact that, in virtue of her passivism, ***her real "causal" mechanisms can't produce, generate or effect anything***, reduces the functional distance between her view and the view of the anti-realist about mechanisms. It is important to appreciate this point. If one wants to distance oneself from those who do not believe in real causal mechanisms, one will have to defend the existence of something that is more efficacious than real mechanisms that can't actually bring anything about. Second, it's not just that a causal mechanism that can't do

[11] Ruth Groff, "Getting Past Hume in the Philosophy of Social Science," in (eds., McKay, Russo and Williamson) *Causality in the Sciences*, Oxford University Press, 2011; pp. 296-316.

anything is functionally equivalent to a causal mechanism that doesn't really exist. It's also not clear why anyone should (or would) characterize such a phenomenon as a causal mechanism. There are many apposite terms for causally inert phenomena. "Causal mechanism," by contrast, would seem to be a misnomer.

The position that I have been describing – realism about causal mechanisms combined with anti-realism about causation – itself comes in various forms. A good example of the type of contradictory stance in question is the account offered by Peter Hedstrom and Petri Ylikoski in a recent piece meant to be "a [critical] review [of] the most important philosophical and social science contributions to the mechanisms approach."[12] Hedstrom and Ylikoski distinguish causal mechanisms from explanations that refer to mechanisms (causal mechanism schemes), and claim for themselves, albeit indirectly, the view that causal mechanisms are the "entities and their properties, activities, and relations" that "produce the effect in question."[13] With respect to sociological effects, they add that it is part and parcel of adopting a mechanisms approach to further stipulate that the entities that produce such effects are individuals rather than macro-level phenomena. Hedstrom and Ylikoski suggest that their depiction of what it is to be a causal mechanism is general enough to encompasses alternate definitions given by a range of thinkers, including Stuart Glennan; Peter Machamer, Lindley Darden, and Carl Craver; Jon Elster; Dan Little; and James Woodward.

While it is the contradiction between their realism about causal mechanisms and their anti-realism about causation that is of primary interest (and I shall turn to it momentarily), philosophical attention to their account reveals another problem, too. The issue is this: as noted, Hedstrom and Ylikoski say that the mechanisms "that produce the effects in question" cannot be macro-level, sociological phenomena. But if macro-level phenomena are not – and cannot be – what produces the effects in question, then we run into trouble. The difficulty is that Hedstrom and Ylikoski also want to say that causal mechanisms are what occurs inside the

[12] Peter Hedstrom and Petri Ylikoski, "Causal Mechanisms in the Social Sciences," *Annual Review of Sociology*, 36: 49-67, 2010; p. 51.
[13] *Ibid.*

so-called black box whereby given macro-level sociological phenomena are related causally. But how can x and y stand in a ***causal*** relation if x's, in principle, cannot cause y's – if, instead, it is mechanism z that produces y's, precisely because x's can't? X is either causally efficacious or it is not. I suspect that some version of this conundrum will show up for anyone who conceives of causal mechanisms as the "how" whereby a designated cause produces a given effect. There will always be the problem of which cause is actually the cause. Is it the cause-cause? Or is it the mechanism-whereby-cause?

The following, meanwhile, is what we get from Hedstrom and Ylikoski with respect to causation itself. "The metaphysics of causation is still hotly debated among philosophers, so it is an advantage that the mechanism-based account of explanation is not wedded to a specific theory of causation. However, the mechanism perspective sets some important constraints for an acceptable theory of causation."[14] Regularity theories can be ruled out, they say, because "whether *a* is a cause of *b* depends on facts about spatiotemporally restricted causal processes, not on what would happen in other similar situations."[15] They also reject any effort to invert the conceptual order by defining causation in ***terms*** of mechanisms, on the grounds that such a definition would be circular and/or would bottom out in a regress at the level of fundamental physics. The theory of causation that is a "more natural complement to the mechanism-based approach," they conclude, is Woodward's.[16] Causation is interventionist-style counterfactual dependency. They then round out their analysis by adding that "[a] mechanism tells us why the counterfactual dependency holds and ties the *relata* of the counterfactual to the knowledge about entities and relations underlying it."[17] This last statement might lead one to wonder if that Hedstrom and Ylikoski are just confused about what they think causation is. Really, it would seem, they don't think that causation is counterfactual dependency at all; really they

[14] *Ibid.*, p. 53.
[15] *Ibid.*
[16] *Ibid.*, pp. 53-54.
[17] *Ibid.*, p. 54.

think that it's the expression of the powers of things. But this can't be, since they specifically reject the idea that causation can be defined in terms of causal mechanisms.

I have introduced Hedstrom and Ylikoski's into the discussion because they recommend defining the causal part of a real causal mechanism as the relation of counterfactual dependence, and I have said that realism about mechanisms plus anti-realism about causation doesn't add up. But their reasoning is problematic at each step along the way. First, in rejecting Humean regularity theory, they conflate the question of what causation is with the question of how one would test for a causal relationship in any given case. While the points are related, the Humean view under consideration is not a methodological edict to the effect that *a* ought to be deemed to be the cause of *b* this time either because or only because it has caused *b* in the past. Rather, it is definitional. The claim from the Humean is that when Hedstrom and Ylikoski say that mechanisms "produce" given effects, the only thing that they can actually mean by "produce" is that a given antecedent mechanism is constantly conjoined with a given consequent. Second, insofar as they endorse the idea that causation is counterfactual dependency, there are obvious complications as soon as we ask what grounds the dependence. The only option available to the passivist is to fall back upon an assertion of order, i.e. regularity across real or logically possible alternate worlds. But Hedstrom and Ylikoski have already rejected regularity theory – albeit on the basis of an arguably ill-formed complaint. As it happens, they have implicitly proposed an alternate grounding for the modal purchase of the relation of counterfactual dependence, viz. "the mechanism [that] tells us why the counterfactual dependence holds." Unfortunately, however, this is the position that they reject when they dismiss the attempt to define causation in terms of causal mechanisms. Finally, the mere fact that they seem to think that counterfactual dependence itself has to be explained suggests that they do not, actually, in the end, think that counterfactual dependence is what causation is.

Still, they say they do. They say that counterfactual dependence is the natural complement to a mechanisms approach. And it is an idea that some find appealing. I have

therefore granted Hedstrom and Ylikoski their professed anti-realism about causation. The proper response, I think, is to reiterate the point that I made earlier: a causal mechanism the "productive" capacity of which amounts to being that which necessarily comes first, is not much of mechanism. And there is no way to get around this. If one rejects a powers-based account of causation in favor of counterfactual dependence, then one is not entitled to causes that can do anything. To be sure, Hedstrom and Ylikoski are not alone in arguing for ostensibly real causal mechanisms that are really just statements of counterfactual dependence that are themselves grounded in an assertion of order. Stuart Glennan, for instance, is at pains to distinguish between mechanisms as he conceives of them and mechanisms conceived as sequences or webs of event. Mechanisms, he says in "Rethinking Mechanistic Explanation," are things – complex systems of interacting parts that produce outcomes via said interaction.[18] But then, when we read the fine print, it turns out that interaction (and presumably the producing that occurs thereby) is to be understood in terms of the semantics of counterfactuals, as he puts it.

(ii) Causes, Causation and Other Confusions

Alongside authors who combine talk of causal mechanisms with anti-realism about causation are those who whose theories display other sorts of conceptual confusion as well. Consider, for example, Neil Gross' "A Pragmatist Theory of Social Mechanisms."[19] Gross begins by reviewing the sociology and philosophy of social science literature on causal mechanisms, then goes on to offer his own account – which, as he explains, differs from alternatives with respect to the substantive psychological-sociological claims that are embedded in it. Gross announces that he is not interested in "social causality as a philosophical concept,"

[18] Stuart Glennan, "Rethinking Mechanistic Explanation," *Philosophy of Science*, 69, September 2002, pp. S342-S353.
[19] Neil Gross, "A Pragmatist Theory of Social Mechanisms," *American Sociological Review*, Vol. 74, June 2009, pp. 358-379.

but rather in "causal processes in the realm of the social."[20] His thesis with respect to the latter is that what causes sociological effects is the actions of individuals or collectivities, specifically their "habitual" responses to "problem situations." These he terms social mechanisms, rather than causal mechanisms, but he presents his view as modifying the literature on causal mechanisms. Gross spends much of the article detailing what he means by "habitual," emphasizing that he – unlike Hedstrom, in his view – does not reduce meaning to belief.[21]

How should we parse such an intervention? Gross is right, it seems to me, that his paper is not about ***causation*** at all. Indeed, nowhere does he tell us what he takes causation to be. The paper is about what Gross believes to be the cause of the class of effects that concern him, viz. sociological effects. In this sense, then, Gross is indeed doing substantive social science, and not philosophy. As always, however, things are not as simple as they may seem. One complication is this: it is hard to imagine how one could pick out either a causal process ***or*** a cause absent any notion of what a specifically causal process or a cause ***is***. And, of course, any notion of what it is to be a causal process (or a cause) will be associated in turn with a conception, if only implicit, of what ***causation*** is. Plainly, one cannot coherently say: "I have no idea what I mean by 'cause,' but I am here to tell you that process p is a causal one; and also that such processes consist of x's, which do the causing when they *phi*." When one makes a causal claim (be it token or type), one thereby votes with one's feet, as it were, with respect to the question of what causation is. Necessarily, that is, one has ***some*** idea of what one is doing in making such a claim. Nor will it do to say: "Well, I know what ***I*** mean by 'cause' (though I'm not telling), but it doesn't matter since nothing hangs on it." This because something does hang on it. It is in the nature of the case that if Humeanism is true, then anything that regularly precedes something else (whether the regularity is direct or via counterfactual dependence) – anything that regularly comes first is, in virtue thereof, a cause. If, by contrast, one assumes causation to be productive, then one will be both perfectly able, and quite likely, to distinguish

[20] *Ibid.*, p. 364.
[21] Gross cites Peter Hedstrom, *Dissecting the Social: On the Principles of Analytic Sociology* (Cambridge: Cambridge University Press, 2005) in particular.

between a causal process and what one will take to be a mere correlation. Thus, while it is important not to equate the question of what causation is with the question of what causes what, it is equally important to see that they cannot be completely separated, either.

But it's not just that in talking about causal processes Gross has necessarily committed himself to ***some*** philosophical definition of causation. We can go a step further. Gross tells us that the type of causal mechanism that is of interest to him is one in which "chains or aggregations of actors" respond, confront and mobilize, in response to "problem situations," thereby causing sociological phenomena.[22] "Respond, " "confront" and "mobilize" are all action terms. The locution indicates that what Gross wants to say is not that there is a constant conjunction between certain intentional states of actors and certain subsequent states of sociological affairs. Rather, the idea is that, in the causing outcomes, the actors in question are actually ***doing*** something. Gross tells us a fair amount about the nature of the doing, e.g., that it is "habitual," that he means by the term "habitual" to refer to responses that are novel as well as to those that are customary or expected (as Hume would have said) and that the potential for novelty is relevant. Thus, in specifying what he takes the causes of sociological phenomena to be, Gross doesn't just endorse any old philosophical theory of causation. He endorses, albeit tacitly, a philosophical theory of causation according to which causes actively produce outcomes (rather than merely preceding them). The insistence on novelty also arguably commits him to free will, though I won't pursue that point here.

As should be clear by now, it is a mistake to think that one can mind one's own empirical business, as a social scientist, and confine oneself to identifying what causes certain kinds of outcome without having an opinion as to what causation is. But there are other problems here as well. For instance, at the same time that Gross sees himself as doing sociology rather than philosophy, he also seems to think that identifying what causes sociological outcomes amounts to stating what a causal mechanism ***is***. But this is to conflate levels of abstraction. If I have told you what causes x-type phenomenon, I have identified the cause of

[22] *Op. Cit.*, Gross, "A Pragmatist Theory of Social Mechanisms," p. 368.

x's. I have not thereby specified (a) what it is to ***be*** a cause, (b) what causation is, or (c) what a causal process is. This is so even if my explanation necessarily relies upon concepts of each, which it will do. Gross also lumps together in a single provisional definition of the term "social mechanism" those accounts according to which the mechanisms don't really exist (the first form of anti-realism discussed earlier) with realist accounts, according to which they do. Further, albeit perhaps rendering the grouping more tenable, he conflates the dispute over the reality of mechanisms with the issue of observability. In contrast to what he calls "realist models," he rightly cites Gudmund Hernes' claim that "a *mechanism* is an intellectual construct that is part of the phantom world which may mimic real life with abstract actors that impersonate humans and cast them in conceptual conditions that emulate actual circumstances." However, he then characterizes the "point of contention" between proponents of Hernes-style "analytic models" and realist models as being an epistemological one, positioning Barbara Reskin's concern with observability as the dichotomous alternative to Gudmund's overt anti-realism. The contested ontological status of posited phenomena is an ontological disagreement, not an epistemic one. For present purposes, however, the point is that by the time that Gross formulates his generic summary account, the issue of the existence or non-existence of the entities that do the causing has disappeared altogether.[23]

Gross' paper is not unique in having problems that philosophy can help resolve. In "Causal Mechanisms: Yes, But …,"[24] for example, John Gerring distinguishes between causal mechanisms and covariational relationships, but then says that causal mechanisms just ***are*** covariational relationships. The initial claim is that mechanisms are the "pathway(s) lying between"[25] covariant cause x and effect y; they tell us ***why*** x causes y (others say "how";

---

[23] "Taken together," he writes, "these considerations suggest the following definition: *A social mechanism is a more or less general sequence of set of social events or processes analyzed at a lower order of complexity or aggregation by which – in certain circumstances – some cause X tends to bring about some effect Y in the realm of human social relations. This sequence or set may or may not be analytically reducible to the actions of individuals who enact it, may underwrite formal or substantive causal processes, and may be observed, unobserved, or in principle unobservable*." (Emphasis in the original.) *Ibid.*, p. 364.

[24] John Gerring, "Causal Mechanisms: Yes, But …" *Comparative Political Studies*, 43 (11), 2010 (Originally published online July 28, 2010), pp. 1499-1526.

[25] *Ibid.*, p. 1506.

Gerring has it as "why"). There are covariational relationships and there are "the causal mechanisms at work in these relationships." It turns out, however, that mechanisms and covariations are not different at all: mechanisms are just covariations that are "intractable."[26] This is bad not just because there is a contradiction involved in saying that mechanisms are and are not the same thing as covariations, but because mechanisms that are just covariations are mechanisms that, manifestly, cannot do anything. Gerring may reply that he never said that they could. All he said is that they are the pathways between a designated cause and an effect, usually a "distal" one, as he puts it.[27] But if pathways themselves are just chains of covariation, then it is very hard to see how they can answer the "Why?" question left unanswered by less mediated covariations. Indeed, it looks very much as though there is only one phenomenon, not two.

But suppose that we grant the operation of not one but two different kinds of causal phenomena, viz., covariations and the pathways that, despite being covariations themselves, account for other covariations. As soon as we grant this we wind up with a regress. Insofar as the mechanisms are themselves covariations, they will require pathways too. Gerring recognizes this. As he would have it, the solution is to say that at a certain point the explanatory distance between the variables becomes so small or self-evident that there is no need to identify a connecting pathway. Gerring thinks that the situation is paradoxical for proponents of causal mechanisms in that, once there is no longer a need for a pathway, the researcher can – and, he thinks, is likely to – declare the final, maximally direct mechanism to be a covariational finding in its own right: one so obvious that it needn't be bolstered by an account of the mechanisms that account for it. Paradoxical or not, the solution is not a genuine solution. The regress occurs because, according to Gerring, there is a follow-up question that the covariational causal claim "X causes y" simply cannot answer. Covariation tells us what, he says; mechanisms tell us why. The final covariation that supposedly stops the regress cannot

[26] *Ibid.*, p. 1517.
[27] *Ibid.*, p. 1506.

tell us ***why*** x causes y any more than could the one that set it in motion. Taking a cue from Gross, we might respond by saying that the philosophers in the crowd are free to keep on worrying about loose ends at the level of logic, but if the final covariation really is self-evident, then in practice there is no regress. Gerring himself seems most interested in the fact that it appears as though even explanations that invoke causal mechanisms do terminate in self-evident covariation, lending credence to the thought, to which he is sympathetic, that there is no real difference between the approach taken by those who believe in mechanisms and the position of those who do not. The reply to both of these lines of thinking is that there is an actual problem here, one that cannot be made to go away merely by declaring that one has lost interest in it. If mechanisms are that which, as a matter of how the world is, account for how or why covariation occurs, then it can't be that they aren't in place when we don't happen to need them, i.e., when the answer to the "How?" or "Why?" question is obvious to us. Either there is something in the world that grounds causally salient covariation, or there isn't.

A different paper by Gerring provides another example of the kind of philosophical confusion that has made its way into the debate over causal mechanisms. In a review article entitled "The Mechanismic Worldview: Thinking Inside the Box," Gerring undertakes to clarify the literature by first reviewing nine definitions of the concept, then plumping for one of them.[28] It is a laudable effort. But what is striking about Gerring's analysis is the way in which it papers over the difference between a powers-based approach and a passivist approach. That it does so is unfortunate enough on its own, given Gerring's claim that "social scientists are now thinking hard about what causation means." (162) It is all the more troublesome considering his further claim that "[f]acing no organized resistance, the mechanismic view" (which Gerring contrasts with Humeanism covariational theories), "now dominates discussions of causality." Far from Humean anti-realism about causation no longer being on the table, Gerring's review of the fault lines in the literature reveals the discussion to consist of a set of disputes ***amongst*** Humeans (or,

[28] John Gerring, "Review Article: The Mechanismic Worldview: Thinking Inside the Box," *British Journal of Political Science*, 38, January 2008, pp. 161-179.

in some instances, disputes defined from a Humean perspective). I cannot put too fine a point on this: ***it is causation conceived as genuinely productive that is off the table, not Humeanism***. Adding to the confusion, Gerring tells readers (a) that Hume, unlike probabilistic thinkers, assumed covariational relationships to be deterministic, but neglects to add that by "necessity" Hume meant a contingent, customary expectation; and (b) that counterfactual arguments are "mechanismic," to use Gerring's term, rather than covariational – the implication being that even Lewis (whom Gerring references in this regard), was somehow an anti-passivist powers theorist, rather than a preeminent contemporary Humean.

Gerring summarizes the nine positions that he considers as follows:

> It is discovered that 'mechanism' has at least nine distinct meanings as the term is used within contemporary social science: (1) the pathway or process by which an effect is produced; (2) an unobservable causal factor; (3) an easy-to-observe causal factor; (4) a context-dependent (bounded) explanation; (5) a universal (or at least highly general) explanation; (6) an explanation that presumes highly contingent phenomena; (7) an explanation built on phenomena that exhibit lawlike regularities; (8) a distinct technique of analysis (based on qualitative, case study or process-tracing evidence); or (9) a micro-level explanation for a causal phenomenon.[29]

(1) is the definition upon which everyone actually agrees, Gerring says; the others are supplemental. I have addressed Gerring's version of (1) in the discussion above, and I will come back to the idea of a pathway or a process "whereby," in the final section of the paper. For now, let me just point out a few things about positions (2)-(9). I do want to note, though, that pathways, explanations and techniques of analysis are very different types of thing. Strictly speaking, it would seem, a mechanism can only be one and not the others.

But the real problem with Gerring's mapping of the conceptual terrain is that it misses the profound disagreement between those who believe that causation is productive and those

[29] *Ibid.*, (Abstract), p. 161.

who reject that view. Definitions (2) and (3), for instance, refer to the criterion of observability. Observability is at issue when it comes to thinking hard about causation because empiricism won out as the dominant philosophy of science As noted earlier, Hume held that all meaningful concepts are grounded in impressions, and that (for better of for worse) there is simply no impressions of a causal power to be had. Present-day empiricist norms come to us through the filter of logical positivism, wherein observation statements came to replace impressions, and certainly productive causation is not the only purported phenomenon the reality of which and/or meaningful talk about which Humeans deny. But this point is this: as Gerring presents the debate over observability, the idea is that some thinkers believe that ***causes*** are or must be observable, while others think that they must be unobservable. The description is accurate, but it is not perspicacious. Crucially, both parties to the debate may very well accept an anti-realist account of causation itself (and generally do), precisely because they agree with Hume that causal powers are not observable.

Definitions (4) – (9) similarly bypass (a) the issue of what causation is, and (b) how competing answers to that question will bear upon the considerations picked out in the various definitions of a causal mechanism. Criteria such as contingency (item 6) versus law-like regularity (item 7), for example, will be understood in completely different ways depending upon whether one holds a passivist or a powers-based metaphysics. For the passivist, causation just ***is*** the fact of order, either in this world (contingent regularity, either at the level of particulars or at the level of universals) or in all logically possible worlds (counterfactual dependency). So-called probabilistic causation is, for this very reason, construed as less causally dense, less deterministic, than non-probabilistic causation is or would be. Indeed, probabilistic non-productive causation is weak enough that Humeans who are incompatibilists think that it allows for the existence of uncaused acts of libertarian free will. Instances of singular causation, meanwhile – where there is no condition of order since the event is unique – are notoriously difficult to conceptualize from this perspective. For the powers theorist, by contrast, causation is causation is causation – viz., the expression of things' ("things'") powers-

to-do. Some instances of doing are law-like and some are not, but this does not tell us that there are different degrees or intensities of causation (deterministic, probabilistic and singular); it tells us only that things of different kinds behave differently. Here too, then, Gerring's overview papers over the divide between realism and anti-realism about causation. Above all, the variable in question – law-like regularity versus unique (and therefore presumptively contingent) occurrence – is not especially salient, for the powers theorist. Whether she is describing types of doing or token instances of doing, the anti-passivist will be interested in the nature of the power(s) involved, not in the fact of regularity. Again, some things have powers in virtue of which they behave in law-like ways, *c.p.*; other things have powers in virtue of which they do not. The important question is what the powers of given things are, and what kinds of processes are constituted, in part or whole, by their display. There is a real sense, therefore, in which definitions (6) and (7) only have significant traction in a Humean environment. This said, the insistence on the part of some that causal mechanisms are singular phenomena may well signal dissonance, vis-à-vis Humean presumptions. Unfortunately the piece does not cut deep enough philosophically for us to know.

Finally, let me say a quick word here about James Mahoney's position. In a paper presented at the American Political Science Association annual meeting in 2003, Mahoney defines a causal mechanism as a hypothetical, necessarily unobservable, productive "ultimate cause," which is believed to be real by the researcher who posits it.[30] Such causes account for dependency relations between observable phenomena, Mahoney says. Indeed, in Mahoney's view, causal mechanisms necessitate, rather than merely raise the likely occurrence of, those patterns of association that they ground. Three aspects of Mahoney's view invite comment. First, Mahoney seems to think that there is a principled reason why causal mechanisms as he

[30] James Mahoney, "Tentative Answers to Questions About Causal Mechanisms," American Political Science Association, annual meeting, Philadelphia, PA, 2003; http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCkQFjAA&url=http%3A%2F%2Fciece.com.ar%2Fciece%2Fwp-content%2Fuploads%2FMahoney-James-Tentative-Answers-to-Questions-about-Causal-Mechanisms.pdf&ei=6RPHUrDWCKWZ2QXYhIC4CQ&usg=AFQjCNEes2KNoekbuySIvKdndgpUXQLYcg&bvm=bv.58187178,d.b2I See also Mahoney, (Review Essay) "Beyond Correlational Analysis: Recent Innovations in Theory and Method," *Sociological Forum*, Vol. 16, Number 3 (September 2001), pp. 575-593.

defines them would (i.e., do) necessitate their effects. The claim is easily defensible if Mahoney means only to say that it is in the nature of patriarchy, say, to yield gender as a significant independent variable. The idea would be that if the causal mechanism in question is patriarchy, then necessarily gender is the factor that is causally salient. But to say this is just to say that patriarchy is necessarily patriarchy. It is what Stephen Mumford and Rani Lill Anjum, referencing Kant, call "categorical" necessity.[31] Categorical necessity does not imply causal necessitation. For example, it does not follow from the fact that a society is patriarchal that gender will in fact determine every particular outcome. Now, given that in Mahoney's model causal mechanisms underwrite dependency relations, and not particular outcomes, we might read him as claiming only that mechanism A causally (and not just categorically) necessitates the dependency relation <B, C>, but then stipulate that the dependency relation ***itself*** is only probabilistic. (Here's what Mahoney says exactly, in a different piece: "if the mechanism actually operates, it will always produce the outcome of interest."[32]) Such a formulation would be fine, I think. But one would have to be very careful not to confuse such a claim with one of causal necessitation at the level of actual events, which is what Mahoney sounds to be endorsing. Next, Mahoney argues that it is in the nature of the case that a causal mechanism cannot be observable. If it were observable, he says, it would have to itself be the effect of some prior cause, and therefore it would not be "ultimate." Here there are two things to be said. First, an originating, productive cause either is or is not such a cause; it can't be both. Of course, we might turn out to have been mistaken in thinking that A is what produces <B,C>. And, anticipating this possibility, Mahoney might mean for us to understand not that a causal mechanism ***is*** an originating cause, but rather that a causal mechanism is that which is ***believed***, at any one moment, to be an originating cause. But – and this is the second thing to be said – there is no reason to think that the observability or non-observability of A would affect whether we ought to believe that it is the originating cause of <B, C>. This because there is no reason to

[31] *Op. Cit.*, Mumford and Anjum, pp. 163-164.
[32] *Op. Cit.*, Mahoney, "Beyond Correlational Analysis: Recent Innovations in Theory and Method," p. 580.

think that there is any relationship between observability and ***being*** an originating or ultimate cause. Mahoney's premise to the contrary is a relic of the passivist picture that he otherwise rejects, the version in which causation is a rubric of order, specifically an unbroken, temporally tensed, deterministic chain of observable causes. His thought seems to be that were a causal mechanism to be (or become) observable, it would thereby exist (or have come to exist) in some more robust way than it had before, such that it too would be necessarily linked to a prior determining cause. The point here is not that Mahoney can't revert back to nomological passivism in the end, if he so chooses, only that elsewhere he endorses productive causation, and a powers-based ontology is an ontology of things with powers-to-do, not one of chains. Third, just as a genuinely ultimate cause can't be both ultimate and not-ultimate, a real causal mechanism cannot be both real and not real. Real-but-unobservable is fine. But Mahoney wants Real-and-yet-only-posited (and thereby not-itself-caused).

The literature on causal mechanisms is full of accounts that don't quite add up in one way or another. The sampling that I have presented is not meant to be exhaustive (or particularly egregious), merely illustrative. I want to turn now to a final set of theories, ones that are consistently anti-passivist, such that the appeal to "mechanism" as a metaphor of productive causation is entirely justified. At a certain level, I want to say, these theories ***do*** add up. All the same, there is further philosophical work to be done in this context too.

(iii) What Exactly is a Causal Mechanism, Again?

Some realists about causal mechanisms also reject passivism. Such thinkers are realists about causal mechanisms but also about causation itself. Roy Bhaskar, joined in short order by Nancy Cartwright, was an early such figure amongst philosophers of social science. Subsequent anti-passivist theorists of social science who have written about what a causal mechanism is include: Doug Porpora, Tony Lawson, Dave Elder-Vass, Phil Gorski and George Steinmetz. I would also include Dan Little and Peter Machamer, Lindley Darden and Carl Craver on this list. There are appreciable differences in the views of these writers, and Little

and Machamer, Darden and Craver are arguably outliers. But all (I think even Little and Machamer, Darden and Craver) believe that causation is productive all the way down. Let's call this a powers-based approach to causal mechanisms. The relevant literature features five main models of a productive or powers-based causal mechanism. These are: (1) a mechanism is the underlying, productive cause of a manifest regularity at the level of events; (2) a mechanism is a power (of a certain kind, perhaps of borne by a certain type of bearer); (3) a mechanism is a series of steps whereby A productively causes B; (4) a mechanism is a process, whereby (or wherein) A productively causes B; (5) a mechanism is a pathway between productive cause A and effect B. What I'd like to suggest by way of conclusion is that regardless of which model one prefers, the idea of a causal mechanism may ultimately be more trouble than it's worth. Even in the hands of anti-passivists, it seems to me, "causal mechanism" is either a misnomer for "cause" or for "power," or it is a gerrymandered category, referring to some mix of causes, powers and processes, or to a sequence of causes plus an assumption about the transitivity of causation. At the same time, though, even if it may not, in the end, cut the relevant object domain at its joints, the notion of a causal mechanism does give us a clue, I think, to the requisite categorical infrastructure of a powers-based account of causal phenomena.

Some powers-based models of a causal mechanism are simply the concept of a cause, renamed. At one end of the continuum in this regard is an approach such as that defended Roy Bhaskar in *A Realist Theory of Science*. Doug Porpora might be seen as endorsing this same view in "Do Realists Run Regressions?" as might Nancy Cartwright in her notion of a "capacity" – and perhaps Mahoney, too.[33] From this perspective (viz., model (1) above), a causal mechanism is that which gives rise to regularity. Here we would say – as Mahoney does explicitly – that causal mechanism A is what produces the ordered or covariational pair <B, C>. Racism, for example (though Mahoney's own examples are more general), is what causes there to be a significant correlation between being black and being poor. It is not, on this model, that being

[33] Douglas V. Porpora, "Do Realists Run Regressions?" in Jose Lopez and Garry Potter (eds.), *After Postmodernism: An Introduction to Critical Realism* (London: Continuum, 2001), pp. 260-266.

black causes one to be poor, ***via racism*** (B causes C, via A). Nor would the claim be that racism is the cause of black poverty, and that being black is the mechanism whereby it does so. (A causes C, via B). Instead, the claim is that racism causes income to be stratified by race (A [productively] causes the [non-productive] covariation between B and C). It is easy to see why those who take this view might be inclined to think that causal mechanisms at the level of the social are by definition structures rather than individuals. Be this as it may, what matters is that the underlying causal mechanism (whatever type of "thing" it might be) just ***is*** the productive cause, from this perspective. The antecedent variable in the produced ordered pair <B, C> is ***not*** the cause. Causal mechanisms are interesting causes, on this version of the mechanisms-as-causes model, in that what they cause are regularities. But they are causes all the same. On the other end of the continuum, in that its proponents are more often passivists than powers theorists, is model (5) above, according to which a mechanism is a pathway between cause A and effect B. Jon Elster is a paradigm example of a passivist who holds this view.[34] We can at least imagine a powers-based version, however, in which each of the steps along the path would be a productive cause on its own, with the final step being the proximate, actual cause of B. Dan Little has said something like this in the past, although his anti-passivism, if I am right to attribute a productive view of causation to him, was ambiguous at the time.[35] His more recent statements sound exactly this way, although sometimes he inserts a "by which" into the formulation, such that it is ***A*** that causes B, not the final step in the pathway that causes B.[36] Given the radical difference between passivist and powers-based understandings of causation, it is important not to conflate passivist and powers-based accounts, be they "pathway" accounts or "steps by which" accounts. Note too that a powers-based ***process*** will not be thought to reduce to a series of steps. I will turn to the powers-based "by which" models - i.e., (3) and (4)

[34] See, e.g., Jon Elster, *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences* (New York: Cambridge University Press, 2007).
[35] See, e.g., Daniel Little, *Varieties of Social Explanation* (Boulder: Westview Press, 1991), chapter 2.
[36] For a more pronouncedly anti-passivist version of the pathway account see most recently Dan Little, *Understanding Society* blog post, October 29, 2013 http://understandingsociety.blogspot.com/2013_10_01_archive.html.

above - in a moment. The point here is that (1) and (5) give us causal mechanisms whose formal character, it seems to me, is fully captured by the term "cause."

A good example of (2) meanwhile, is Phil Gorski's ECPRES model.[37] Causal mechanisms, Gorski says, are "*emergent causal powers of related entities within a system*."[38] [Emphasis in the original.] The formulation contains an "of," such that we might think that it combines the notion of a power with the notion of a cause, or even that here too a causal mechanism just ***is*** a cause (one with certain kinds of powers, given the kind of "thing" that it is, as is the case with all causes). There is no harm in parsing the definition this way instead. If we do, then it can be grouped with (1) and (5). Gorski sometimes talk this way himself: "In thinking about what kinds of causal powers we should and should not ascribe to which mechanisms …"[39] makes it sound as though mechanisms are powers-bearers, not powers. Still, Gorski's focus is on the powers part, it seems to me. He is not alone in this respect. Bhaskar, for example, while casting mechanisms as underlying causes, seems at the same time to identify them with powers, inasmuch as causal mechanisms are assigned to what Bhaskar calls the domain of the Real, which is comprised of powers. Similarly, in the context of contemporary analytic metaphysics, Stephen Mumford and Rani Lill Anjum direct their attention almost exclusively to powers, rather than to the bearers of powers. Nothing hinges upon my reading of Gorski's position, though. The point is simply that if we do take the ECPRES definition at face value, we have causal mechanisms that are equivalent to powers. As with causal mechanisms that are equivalent to causes, we might begin to wonder if there is a need for the category of powers under a different name.

Models (3) and (4) are a little different from (1), (2) and (5). As a preliminary, let's register that from a Humean perspective a process is nothing other than a sequence of facts, or static states of affairs. For a Humean, then, there is no difference between (3) and (4); in

---

[37] Philip Gorski, "Social 'Mechanisms' and Comparative-Historical Sociology: A Critical Realist Proposal," in (eds.) Peter Hedstrom and Bjorn Wittrock, *Frontiers of Sociology (Annals of the International Institute of Sociology, Volume II* (Brill: Leiden, 2009), pp. 147-194.
[38] *Ibid.*, p. 147.
[39] *Ibid.*, p. 158.

referring to steps as a "process," she will say, proponents of (4) have simply helped themselves to a metaphor. For the non-Humean, by contrast, a process is dynamic, while a sequence of steps is not dynamic. This will be the case even when the sequence is conceived as being a chain of ***productive*** causes, rather than a sequence of facts or static states of affairs. A "process," in turn, will be conceptualized as something other than a chain of discrete productive causes. It is the powers-based versions of (3) and (4) that I described as gerrymandered. In the case of (3), I say have characterized it that way because the concept of a causal mechanism is serving as an umbrella for (i) the concept of a productive cause; (ii) the stipulation that productive causation, like non-productive causation, is transitive; (iii) a further assumption that the transitivity in question is local, i.e., that it holds between A and B, but not prior to A, thereby allowing for A to itself be an originating productive cause; (iv) a final assumption, at least on the part of some proponents of this model, that the reason why A doesn't cause B directly is because A, in virtue of being a macro-level phenomenon, can't. Even if one agrees with (i) – (iv), bundling them into a single concept is not ideal. One reason to hesitate is that the issue of transitivity is far more complicated philosophically if one is not a passivist. If impression (or statement or fact or event) A logically implies B, or is constantly conjoined with B, and B in turn with C, then it follows relatively easily that A logically implies or will be constantly conjoined with C. But this is not so if we assume that productive cause A has certain powers, productive cause B entirely different ones. Just because A has the power to cause B, it does not follow that A can do what B can do, as every parent knows.

A powers-based (4) gives us causal mechanisms defined as "a ***process*** whereby," rather than "a chain of productive causes by which." In this case, it is not that there are too many assumptions packed into one concept, but that the categories of cause and process are not properly combined. The "whereby" flags the problem. It's a trickier locution than it seems. A process, as distinct from a series of discrete productive causes, is unified; it is a single diachronic expression of a particular type of doing. What exactly is it that is being claimed, then, if one says that the temporally extended expression of a particular kind of doing is that "whereby" a

productive cause brings about an effect? The idea would seem to be that the process is an instrument, a means – a means by which (once again) a cause causes an effect. However this doesn't seem right. It will be perfectly fine to say that causes behave in the ways that they potentially can, given their dispositional properties, and we will also want to say that causation itself involves the display of such properties in or as processes of various kinds. But the display of a thing's power to do x is not an instrument with which it does x; it just ***is*** the doing of x. Nor, more generally, is "doing" itself a tool for the causing or bringing about of causation; rather, doing just ***is*** causation.

If models (1), (2) and (5) give us causal mechanisms that are equivalent to causes or powers, and models (3) and (4) give us mechanisms that are either chains of causes plus assumptions about transitivity, or processes that don't stand in quite the correct relation to causes or to causation, the obvious question is whether or not the concept of a causal mechanism should be retained at all. The important thing, it seems to me, is to get one's basic categories right. I suspect that "causal mechanism" may not be one, for the reasons I've said. When I say that it is more trouble than it is worth, what I mean is that we do not need causes, causation ***and*** causal mechanisms. But the concept does play a useful role as a kind of sign-post. What it does above all in this regard, I think, is remind anti-passivists that we may very well need the category of "process" along with those of "cause" and "power" – even if processes are not actually "that ***whereby*** causes do what they have the power to do," but instead are simply the doing itself.[40] The very next question that one would want to ask would be whether or not a category of "process" would be redundant vis-à-vis that of "causation." But it is worth thinking about seriously. In addition, the concept of a causal mechanism draws attention to two notable facts about social reality. The first is that emergent sociological phenomena, assuming that one believes that such "things" exist, have different kinds of powers than do individuals. Social structures can't talk, or pick things up. If one holds that causation is productive, one will have to be able to say how, and/or in what sense, producing is done by

[40] See Brian Ellis, *Scientific Essentialism* (Cambridge: Cambridge University Press, 2001).

such phenomena. Just what is it that macro-level phenomena can do? It won't be enough to stipulate that A causes B "by way of" a different kind of cause, CM, with more familiar powers, since then it's CM and not A that is the cause. This said, the "by which" or "by way of" signals that there is something there to be theorized. Finally, the notion that the causation of B by A often occurs via a chain of causes, rather than directly, suggests that what occurs at the level of the social is complex – more complex than the are causal scenarios that often figure as examples in work on causation by philosophers. It is important an important point, since it suggests that the simpler ontologies illustrated by such examples may be incorrect.

The lesson to be learned from all of this, however, is that there is a need for ***more*** philosophy, rather for less.