An Anti-reductionist Account of Singular Causation

The Monist 92:1 (Jan 2009): 133-152

When it comes to the conceptual analysis of causation, no one theory enjoys widespread acceptance. It is generally acknowledged that counterfactual and probabilistic accounts of causation are subject to serious objections, at least when they are offered as reductive conceptual analyses. Process or transference accounts—again, if offered as reductive conceptual analyses— also face major difficulties. Given the prolonged failure to uncover a successful reductionist conceptual analysis of causation, anti-reductionist accounts can become attractive. Perhaps the reason that we haven't yet found a successful reductionist analysis is simply that there isn't one to be found—perhaps causation is primitive. What is more, an account which posits a primitive relation of causation between particulars can handle many of the cases that cause trouble for reductive accounts. Yet the move toward primitivist or anti-reductionist accounts is resisted by many philosophers. After spelling out his reductive analysis of productive causation, for example, Ned Hall writes:

> I do not mean to pretend…that the analysis I have offered stands in no need of detailed elaboration or defense. Of course it does. But that task can be left for another occasion, since my aim here is quite modest: I mean only to make it at least somewhat plausible that a reductive analysis of production can be had—thereby blocking the objection that once we distinguish production and dependence…we will be stuck with "production" as an unanalyzable causal primitive. That would indeed be unfortunate.[1]

Hall's antipathy to unanalyzable causal primitives is not uncommon. What is less common is the marshalling of arguments backing this sort of antipathy. Arguments against anti-reductionist positions are comparatively rare and typically rather brief.

In this paper, I want to articulate and defend a particular anti-reductionist view. In section one I try to sort out the meanings of several terms that tend to do similar work in the literature: non-reductionism, anti-reductionism, singularism, and primitivism. Doing so will help me sketch (in

---

[1] Hall, "Two Concepts of Causation," in *Causation and Counterfactuals*, eds. J. Collins, N. Hall, and L.A. Paul, pp. 225-276 (Cambridge, MA: The MIT Press, 2004), p. 265.

section two) an anti-reductionist account that can withstand extant objections to anti-

reductionism. In the remaining sections I consider three such objections, one traceable to Hume,

and two due to Hall.

I.   Anti-reductionism, singularism, and their neighbors

The project of assessing non-reductionist accounts of causation is made more difficult by the

presence of several over-lapping but distinct terms employed to contrast various non-reductionist

views from their reductionist rivals. In this section I'll try to sort out the meanings of these terms,

namely: anti-reductionism, non-reductionism, primitivism, and singularism.

John Collins, Ned Hall, and L.A. Paul explain the distinction between reductionist and anti-

reductionist accounts of causation as follows:

> Probably the most useful distinction [amongst accounts of causation] to make at the
> outset is that between accounts that do and accounts that do not attempt to *reduce* causal
> facts to facts about what happens, together with facts about what the laws are that govern
> what happens. (We have a permissive sense of "what happens" in mind: It is to include
> facts about what objects exist where and when, and what categorical properties and
> relations they instantiate.) …We will henceforth label "reductionist" any positions
> according to which causal facts can be reduced to categorical plus nomological facts, and
> label "antireductionist" any position that denies this claim.[2]

Causal facts are facts about what causes what, e.g. the fact that Billy caused the shattering of the

window. Nomological facts are facts expressing the laws of nature. Categorical facts are non-

modal facts about the (non-modal) properties and behavior of things, like the fact that the Statue

of Liberty is on Liberty Island.

Following the last sentence of the Collins *et al*. quotation reproduced above, let us say that,

(R)   A reductionist account of causation is one according to which all causal facts can be
       reduced to categorical plus nomological facts, while an antireductionist account is one
       which denies the claim that all causal facts can be reduced to categorical plus
       nomological facts.

The concept of "reduction" is obviously central here, and we might wonder what exactly it

amounts to. As it turns out, reduction is not just supervenience. Perhaps in some contexts, to say

---

[2] Collins et al., "Counterfactuals and Causation: History, Problems, and Prospects," in *Causation and Counterfactuals*, p. 12.

that a fact F reduces to some set of other facts might merely mean that F supervenes on those other facts. But for Hall, at least, this is not what is meant by "reduction" in the present discussion.[3] As Hall points out, the notion of supervenience does not entail asymmetry, while the relevant notion of reduction does.

If reduction is not supervenience, what is it? To bring the concept into clearer focus, consider an example. Suppose it is a fact (F) that Ken is a bachelor. It seems right to say that there is nothing to this fact over and above the facts that (F1) Ken is a male human being, (F2) Ken is unmarried, and (F3) Ken is of marriageable age. Given that F1, F2, and F3 obtain, nothing more need be posited in reality for it to be true that F obtains. We could even say that facts F1, F2 and F3 make up fact F.

I'll take this example as a paradigm of fact-reduction. To say that one fact is reducible to a set of other facts is to say that the set of other facts is "all there is" to the first fact, which means that if the set of other facts all obtain, then nothing more need be posited in the ontological structure of reality for it to be the case that the first fact obtains. Where fact F is reducible to facts F1-Fn, F1-Fn make up or constitute F, and the obtaining of F *just is* the obtaining of F1-Fn.

The contrast between reductionism and anti-reductionism presupposes a distinction between basic (or primitive) facts and derived facts. Derived facts are facts that are reducible to other facts. Basic facts are facts that are not reducible to other facts. The idea seems to be that some facts are reducible to other more foundational facts, and these in turn either are themselves basic or are reducible to basic facts. (The only other option requires us to posit an infinite regress of fact-reduction.) The core of the disagreement between the anti-reductionist and the reductionist, then, is about which facts are basic. According to the anti-reductionist any causal fact is either basic, or reducible only to sets that include some causal facts. Given the assumption that there are no infinite regresses of fact-reduction, this implies that at least some causal facts are basic.

---

[3] Ned Hall, "Causation," in *The Oxford Handbook of Contemporary Philosophy*, eds. Frank Jackson and Michael Smith (New York and Oxford: Oxford U.P., 2005), p. 508.

According to the reductionist, causal facts are never basic—all causal facts are reducible to non-causal facts. So an anti-reductionist might be comfortable thinking that (F) the fact that electron one causes electron two to move is a basic fact, while a reductionist would have to hold that F is reducible to some set of non-causal facts.

The term "non-reductionism" is sometimes used synonymously with "anti-reductionism."[4] It is, however, worthwhile to distinguish the two, so I will define "anti-reductionism" as in (R), and use "non-reductionism" to describe a theory which does not attempt to reduce causal facts to categorical plus nomological facts. It is one thing to deny that such a reduction can be made (anti-reductionism); it is another to simply not attempt such a reduction (non-reductionism). One can imagine a philosopher offering an account of causation which does not reduce facts about causation to non-causal facts (it might leave a causal primitive in the analyzans), but yet does not go so far as to deny that such a reduction can be made (the question of whether the remaining causal primitive can itself be reduced would be left open). Such an account would be neither reductionist nor anti-reductionist according to (R); it would appropriately be labeled non-reductionist on the current suggestion.

"Primitivism" is sometimes used to name the view that the relation of causation is primitive and unanalyzable. As Jonathan Schaffer has pointed out,[5] we might profitably distinguish between conceptual primitivism and ontological primitivism. One can develop Schaffer's suggestion as follows: a given concept C is *not* conceptually primitive just in case it can be given an adequate conceptual analysis that includes no reference to C (or a concept equivalent to C, or a concept that is itself analyzed in terms of C) in the analyzans. Bachelorhood, for example, is not conceptually primitive, as it can be given this conceptual analysis: X is a bachelor iff (i) X is a male human being, (ii) who is unmarried, and (iii) who is of marriageable age. Conceptual primitivism with respect to causation, then, is a view about our *concept* of causation, namely, the

---

[4] As in Hall, "Causation," p. 510.
[5] At the Causation and Responsibility Roundtable, organized by Michael Moore and Richard Fumerton, 2006.

view that our concept of causation cannot be given an adequate conceptual analysis that includes no reference to causation (or an equivalent concept, or a concept that is itself analyzed in terms of causation) in the analyzans.

By contrast, ontological primitivism with respect to some item or class of items is a view about those items themselves, not a view about our concepts. Say that ontological primitivism with respect to X is the view that X is a fundamental feature of reality, in the sense that X cannot be reduced to some set of more basic entities. Ontological primitivism with respect to causation, then, would be the view that the relation of causation is a basic or primitive relation; it cannot be reduced to any other relation, or set of relations, or any other ontological items whatsoever.

While the notions of anti-reductionism and ontological primitivism cover much the same ground, the notion of primitivism is more fine-grained. Anti-reductionism is a claim about the whole class of causal facts; if even some subset of the class of causal facts is basic, then anti-reductionism is true. Primitivism gives us more precision when classifying theories, for it leaves room to say that one particular causal relation (say, production) is ontologically basic, while some other causal relation (say, causation by absences) is not.

An additional distinction that will be helpful is that between a relation and an instance of a relation. For example, talk about the being-to-the-left-of relation that obtains between object A and object B can be ambiguous between (i) talk about the particular instance of the being-to-the-left-of relation that obtains between A and B, and (ii) talk about the being-to-the-left-of relation itself. When the claim is made that causation is conceptually primitive, that claim is not best taken as a claim about our concept of a particular instance of some relation, but as a claim about our concept of the relation (be it understood in realist or nominalist terms). A claim about ontological primitivism, by contrast, could be made either about an instance or the relation itself. An anti-reductionist partial to the metaphysics of tropes (particularized properties and relations) might hold that each individual instance of some causal relation is ontologically primitive. An

anti-reductionist partial to platonic realism might hold that some causal relation itself is ontologically primitive.

Finally, singularism. A singular causal statement is a statement which expresses a causal relation between particulars (usually, particular *events*), e.g. "John's habit of smoking was a cause of his developing lung cancer." A general causal statement is a statement which expresses a causal relation between relata described more generally, e.g. "Smoking causes cancer."

The phrase "singular causation" is used differently by different authors. In its most basic sense, a relation of singular causation is just a causal relation whose relata are particulars. This is the sense in which I will use "singular causation." It should be noted, however, that "singular causation" can also bear the more determinate meaning of "causation that does not exemplify a pattern captured in a law."[6] Michael Tooley has used the term "singularism" to express this same view: in his usage a singularist view of causation is a view according to which there could be at least one singular causal relation that does not fall under a law.[7] By my lights, this use unnecessarily restricts the meaning of "singularism." A philosopher adopting a dispositionalist account of causation, who holds that laws are just summaries or descriptions of the ways in which the causal powers of things get exercised, might hold that there can be no singular causation that doesn't fall under some law. (If there are any causal powers at all, then there will be some descriptions of how they manifest themselves, and every exercise of a causal power will fall under one of those descriptions…or so it might plausibly be thought.) But that same philosopher would be described by many philosophers as holding a singularist view of causation, insofar as she denies that causal laws are more basic than singular causal relations.

Accordingly, I will define *a singularist theory of causation* as a theory according to which (i) there is a causal relation between particulars, and (ii) this relation is either more basic than, or at

---

[6] Richard Swinburne, "The Argument from Laws of Nature Reassessed," in *Debating Design: From Darwin to DNA*, eds. W.A. Dembski and M. Ruse (New York: Cambridge University Press, 2004), p. 308, n. 17.
[7] Personal correspondence.

least not less basic than, general relations between types of events, like those expressed by laws of nature or regularities of succession. Singularist theories of causation deny that general relations between types of events (like the relation expressed in "Electrons repel other negatively charged particles") are more basic than singular relations between particulars (like the relation expressed in "This electron is repelling that electron"). Singularism, finally, is just the thesis that some singularist theory of causation is true.

II. An anti-reductionist account of singular causation

In this section I try to sketch an anti-reductionist account of causation which can meet the challenges directed at other versions of anti-reductionism. In order to motivate my account, I must first discuss a view we might call naïve anti-reductionism:

(NAR) Where C and E are actually occurring events, anytime we can say that C is a cause of E, there exists between C and E an ontologically primitive instance of the relation of singular causation.

Naïve anti-reductionism faces difficulties because it ignores the difference between direct and indirect causation (more on these terms below—for now: direct causation involves no causal intermediaries, indirect causation does). To see the problem, suppose that John drinks a cup of coffee and that this causes him to perk up. Call the event consisting in John's drinking a cup of coffee "C" and the event consisting in John's perking up "E". Since we can truthfully say that John's drinking a cup of coffee is a cause of his perking up, NAR implies that there is an ontologically primitive instance of the singular causal relation between C and E. But this seems false—the causation between C and E is indirect causation, and can be ontologically reduced to a chain of intermediate causal relations. This is easiest to see if in fact the set of causal intermediaries connecting C and E is finite. In that case, the relation between C and E is reducible to the set consisting of the instance of causation connecting C to the first intermediate cause, the instance of causation connecting the first intermediate cause to the second,…and so forth up to the instance of causation connecting the last intermediate cause to E. And even in a case where

the causal chain between C and E is dense,[8] it seems right to think that the causation between C and E can be ontologically reduced—once we have in place the dense causal chain between C and E, it is unnecessary to posit some ontologically primitive instance of causation which links C and E directly.

So NAR has a problem with cases of indirect causation. To deal with this problem, and others like it, I'll introduce a more nuanced version of anti-reductionism. The account I will propose consists of claims about (i) the analogical character of causation, (ii) the core of the concept of causation, (iii) a paradigm type of causation: production, and (iv) other types of causation, viz., those involving absences. For ease of reference, I'll give this account the name of scholastic anti-reductionism (SAR).[9]

The first essential claim of scholastic anti-reductionism is that the concept of a cause, and hence the concept of causation, is an analogical and somewhat vague concept. (More on this below.) Second, the core of the concept of causation is derivativeness. As Anscombe puts it:

> [C]ausality consists in the derivativeness of an effect from its causes. This is the core, the common feature, of causality in its various kinds. Effects derive from, arise out of, come of, their causes.[10]

Where Anscombe uses "derive from, arise out of, come of", we might also use "come from" – effects come from their causes. Or we could put it in Aquinas's way: a cause is a source of the being of its effect.[11] Given what Aquinas means here by being (*esse*), this is roughly equivalent to the claim that a cause is either the source of the entire existence of the effect, or the source of some aspect of the effect, e.g. the fact that it has some property or feature. The crucial idea here for our purposes is the notion of a *source of being*, which, I take it, is another way to express the

---

[8] To say that the causal chain between C and E is dense is to say that between any two events in the chain, there is another event in the chain.

[9] The account I offer is influenced by Aquinas and Suarez, hence the name.

[10] G. E. M. Anscombe, "Causality and Determination," in *Causation*, eds. Ernest Sosa and Michael Tooley (New York: Oxford University Press, 1993), pp. 91-2.

[11] "Every cause is related as a principle [read: source – MR] to the being of its effect" [*Scriptum super Sententiis*, M.F. Moos, ed. (Paris, 1947), 1.29.1.1c]. And again: " The name 'principle' expresses a relation of origin absolutely, but the name 'cause' expresses a relation of origin with reference to the being of the thing which proceeds from the cause" [*Scriptum super Sententiis* 3.11.1.1 ad 5].

idea that effects come from or derive from their causes. All these glosses on the concept of derivativeness are not meant to conceptually analyze derivativeness in terms of something else—on SAR, the concept of derivativeness is conceptually primitive.

Third, a paradigm type of causation is production, or productive causation. By "production" I mean to pick out a type of causation which Ned Hall has contrasted with mere counterfactual dependence, a type of causation which we evoke "when we say of an event *c* that it helps to *generate* or *bring about* or *produce* another event *e*".[12] Production is close to what Aquinas would have called *per se* efficient causation. The following cases are examples of production: a fire causes water to be heated (or, the fire's being hot causes the water's being heated, if you like); a massive object produces motion in some other object with mass; parents generate a new human being; God creates the universe *ex nihilo*.

Production comes in two varieties: direct and indirect. The concept of production is conceptually primitive. We can define direct and indirect production in terms of production and one additional notion—the notion of causing something *by* causing something else. I take it that we have some grasp of what it means to bring about or produce an effect E by bringing about some distinct effect, which in turn brings about E. For example, I can bring it about that the window breaks by bringing it about that a stone is hurtling through the air towards the window. To put the point in terms of event-causation, the event which is my throwing the stone (C) is a productive cause of (E) the breaking of the window, on account of being a productive cause of (M) the stone's hurtling through the air in the direction of the window. That said, we can define indirect production as follows:

(IP)   An event C is an indirect productive cause of an event E iff
        C is a productive cause of E, on account of being a productive cause of some event M, which is itself a productive cause of E.[13]

Direct production is then defined as:

---

[12] Hall, "Two Concepts of Causation," p. 225.
[13] Note that the definition of indirect production will apply in cases where C and E are connected by a finite causal chain *and* in cases where C and E are connected by a dense causal chain.

(DP)   An event C is a direct productive cause of an event E iff C is a productive cause of E, but not on account of being a productive cause of some event M, which is itself a productive cause of E.

As noted, production is conceptually primitive. But there is more to say than this. While direct production is not conceptually primitive (it can be analyzed in terms of production, as above), any instances of the relation of direct productive causation are ontologically primitive, which is to say that instances of the direct production relation are a fundamental feature of reality; they cannot be reduced to some set of more basic entities. Like direct production, indirect production can be given a conceptual analysis in terms of production. But unlike direct production, indirect production is not ontologically primitive. Where C is an indirect productive cause of E, the relation between C and E can be reduced. If the causal chain between C and E is not dense, the relation between C and E can be reduced to the set of instances of direct production that link the members of the causal chain between C and E. If the causal chain between C and E is dense, the relation between C and E is somehow a matter of the "shorter" relations between C, all the intermediate causes, and E (though exactly how this reduction will go is not entirely perspicuous).

Given all this, it's clear that scholastic anti-reductionism is indeed an anti-reductionist account. For facts about instances of the direct production relation are, of course, causal facts, and SAR denies that such facts can be reduced to categorical plus nomological facts (since SAR denies that such facts can be reduced to anything).

SAR is also a singularist account of causation. Since the instances of the direct production relation are primitive, we can say that the relation of direct causal production is not less basic than general relations between types of events like causal laws or regularities. Add to this the fact that SAR asserts that there are causal relations between particulars, and it follows that SAR counts as a singularist account.

Next, on the account I am presenting, there are many types of causation in addition to direct and indirect production. Suppose, for example, that Jill pulls an oblivious Jack out of the way of an oncoming bus. In some linguistically acceptable sense of the word, Jill's action is a *cause* of Jack's not being hit by the bus. The cause here is a positive event (Jill's action, i.e. the event of Jill's pulling Jack). But, at least at face value, the effect is a so-called negative event (Jack's not being hit by the bus), which is just to say that the effect is an absence of a certain sort of positive event (a Jack's-being-hit-by-the-bus sort of event). Since an absence does not exist, properly speaking, any sense in which an absence is produced is a different sense than the sense in which some positive event is produced in a case of production. So while there is reason to think some type of causation is at play here, there is also reason to think that it is not ordinary production. Call it causation of an absence.

There can also be causation *by* an absence. And there are cases where causation of and by absences get linked together to form various sorts of chains.[14] The claim that causation is an analogical concept allows us to say (with Jonathan Schaffer)[15] that such cases do count as causation. But because the other sorts of causation are in various ways ontologically derivative (as compared to the ontologically basic relation of direct production), we can say that they are not cases of the most metaphysically central type of causation (production).

If causation is analogical, so that there are several different types of causation, we might well ask what unifies the different types, such that they are all types of *causation*. One could go several different ways at this junction; I'll consider two. First, one might develop Anscombe's suggestion that derivativeness "is the core, the common feature, of causality in its various kinds."[16] Perhaps all we need to do, to explain the unity of the concept of causation, is to point

---

[14] For an excellent discussion of the different sorts of causation involving negative events, see the work of Phil Dowe. Dowe himself considers such cases to be cases of "quasi-causation" rather than causation.
[15] Jonathan Schaffer, "Causes need not be Physically Connected to their Effects: The Case for Negative Causation," in Christopher Hitchcock (ed.), *Contemporary Debates in Philosophy of Science* (Malden, MA: Blackwell Publishing, 2004), pp. 197-216.
[16] Anscombe, "Causality and Determination," p. 92.

out that all the types of causation have derivativeness in common. One billiard ball's motion is from, or is due to, the impact of another ball; Jack's insomnia is due to his failure to take his medication; Jack's missing the bus is due to his inattention. While the sense or way in which the effect is due to the cause differs in these cases, there is nevertheless a commonality here, and we evidently have the ability to discern the presence of that commonality.

Or perhaps we should say more, following a suggestion of the late medieval philosopher Francisco Suarez. Discussing the connection between what he calls *per accidens* efficient causes and *per se* efficient causes, Suarez asserts that "a *per accidens* cause is not a true cause but is instead called a cause because of some relation or similarity to a cause or because it is conjoined with a cause."[17] Adapting this thought, and being a bit more permissive about what we allow as a "true" cause, we might say this: types of causation other than production count as causes because of some relation or similarity to production.

How exactly would we then cash out the relations and/or similarities between production and other types of causation? Again, there are several options. To illustrate them, I'll use our example of ordinary causation of an absence: Jill's action (a positive event—call it "A") is a cause of Jack's not being hit by the bus (a negative event—call it "not-B"). Metaphysically, what's going on is that A occurs, and produces a positive event E (say, Jack's being moved to the sidewalk) which in the circumstances is sufficient for the non-occurrence of a B-type event (a Jack's-being-hit-by-the-bus-type event).

A first way to explicate the link between production and ordinary causation of an absence is via their similar connections to counterfactuals.[18] In most cases of production, there are no back-up producing causes waiting in the wings, so that we have this relation between production and counterfactuals:

---

[17] Francisco Suarez, *On Efficient Causality: Metaphysical Disputations 17, 18, and 19*, trans. Alfred J. Freddoso (New Haven, Conn.: Yale University Press, 1994), 17.2.2, p. 11.
[18] I'm helped here by Dowe's work.

> Usually, if C is a producing cause of E, then it follows that if C had not occurred, then E would not have occurred.

In cases like those of Jack, Jill, and the bus, we have a similar relation:

> Usually, if A is a cause of not-B, then it follows that if A had not occurred, then not-B would not have "occurred."[19]

So "C is a producing cause of E" and "A is a cause of not-B" have the same relation to a certain sort of counterfactual. This similarity makes it suitable for us to consider ordinary causation of an absence to be a kind of *causation*. Other similarities relating to counterfactuals could likewise be pointed out to link the other types of causation involving absences to production.

Additional, complementary ways of explicating the similarity between production and causation involving absences are available. We might even use a sort of recipe for linking production and cases of causation involving absences: (i) pick one of what are sometimes called the conceptual connotations of causation (counterfactual dependence, statistical relevance, agential means, grounds for evidence, grounds for explanation, and moral responsibility),[20] (ii) show that production has a distinctive relationship to it, and then (iii) show that a given variety of causation involving absences has a similar relation to it.[21]

I now turn to a defense of SAR. In what follows I examine three objections that have been raised against anti-reductionist or primitivist views. I'll try to show that in each case the objection in question fails as an objection against SAR.

III. The unobservability argument

A first objection, epistemological in nature, is summarized by Jonathan Schaffer:

> [This argument] traces back to Hume and maintains that primitivism conflicts with the existence of causal knowledge. After all, one might argue in a Humean vein, all we can

---

[19] The scare quotes are in place because not-B is a negative event. For not-B to not "occur" is just for some Jack-being-run-over-by-the-bus type event to occur.

[20] This list is from Jonathan Schaffer, "Causes need not be Physically Connected to their Effects: The Case for Negative Causation," p. 199.

[21] I owe this point to Phil Dowe, who has made a similar point (though for a different purpose), in "Causes are Physically Connected to their Effects: Why Preventers and Omissions are not Causes," *Contemporary Debates in Philosophy of Science*, p. 194.

observe is sequences of events; as such, we could never come to know any facts about causal connection if connection is anything over and above such sequences.[22]

To better understand this argument, begin by considering the Humean claim that causation is just regular succession (X is a cause of Y iff Y follows X, and things like Y regularly follow things like X). If causation is just regular succession, it is easy to see how we could have causal knowledge. For we can observe, and therefore know, that Y occurs after X occurs. We can also observe, and therefore know, that things like Y occur after things like X. If causation is just regular succession, knowing those facts is sufficient for knowing that X is a cause of Y.

The objector isn't committed, of course, to the view that causation *is* regular succession—more complicated reductive analyses of causation will afford the same sort of explanation of our ability to attain causal knowledge. We can know facts of the form "X is a cause of Y" because we can observe the sorts of facts to which facts of the form "X is a cause of Y" can be reduced.

In contrast, according to the objector, causal knowledge would elude us if causal facts were not reducible. Presumably, the reason for thinking this has to do with (i) a commitment to the view that our knowledge must be based on what we observe, and (ii) the belief that causal facts are not themselves directly observable. The objector concludes that causation must be ontologically reducible, since we do indeed have causal knowledge. Call this the unobservability argument.

One can see why someone might hold that causal facts are not themselves directly observable. Suppose that A sees B kicking a soccer ball, and sees the soccer ball straightaway move off in the direction B kicked it, and comes to know that B's kick caused the soccer ball to move. The neo-Humean is likely to say that while A does observe, say, the event which consists in B's kicking the ball and the event which consists in the ball's movement, A does not observe the causal *connection* between the two. A sees B's kick, and the ball's movement, but A does not

---

[22] Jonathan Schaffer, "The Metaphysics of Causation", in *The Stanford Encyclopedia of Philosophy (Spring 2003 Edition)*, ed. Edward N. Zalta, section 2.1, URL = <http://plato.stanford.edu/archives/spr2003/entries/causation-metaphysics/>.

see B's kick *causing* the ball's movement. The fact that B's kick is a cause of the ball's movement is not itself contained in A's perceptual experience, in the way that the ball's color, shape, and position are.

But what follows from this? Granting for the sake of argument that causal facts are never directly observable, the neo-Humean's conclusion that 'we cannot come to know causal facts unless they are ontologically reducible to non-causal facts' follows *only if* something like the following is true:

    (1)   We could never come to know any facts of the form "X is a productive cause of Y" unless either (a) such facts were themselves directly observable, or (b) such facts were reducible to facts that are all directly observable.

Yet this epistemological principle has little to recommend it. (1) is an instance of a more general principle:

    (P)   We could never come to know a fact F unless either (a) F were itself directly observable, or (b) F were reducible to facts that are all directly observable.

If (P) had a claim on our belief, then (1) would too. But (P) is false. If (P) were true, we could not know the fact (F1) that our perceptual faculties are generally reliable. The reliability of our perceptual faculties is just not the sort of thing that one can directly perceive. Nor does F1 seem to be reducible to facts that are all directly observable. Perhaps F1 is reducible to the facts that (F1a) our visual faculties are generally reliable and (F1b) our auditory faculties are generally reliable, and so forth. But these sub-facts are not themselves directly observable. One does not directly observe that one's vision, for example, is reliable. One directly observes colors, shapes, motions, and, in a slightly less direct sense, one observes objects. But one does not directly observe the fact that what one seems to see is usually really there, as one seems to see it.[23]

---

[23] One might try to give an argument that one's vision is reliable, based on knowledge derived from direct observation through one or more of the other senses. But such an argument would be circular as an argument that our perceptual faculties are reliable, because it would have to depend on at least one of our perceptual faculties to furnish some of its premises.

If (P) were true, then, we could not come to know that our perceptual faculties are generally reliable. But we do in fact know that our perceptual faculties are generally reliable, so (P) is false. And thus we have no reason to accept (1).

Upon reflection, it becomes apparent that the epistemological assumptions backing the unobservability argument are ungrounded. Both externalists and internalists about knowledge should grant this, as I will now try to show.

The objector advancing the unobservability argument considers a situation where a fact F is knowable, but not itself directly observable. He then concludes that F must be reducible to facts which are directly observable, for otherwise (he thinks) there would be no way to explain how F is knowable. This is a hasty inference. There are other ways we could explain how F is knowable in such a situation.

Suppose a fact G is neither directly observable nor reducible to facts which are. Now suppose I form the belief that G on the occasion of having a certain perceptual experience, and suppose further that the belief-forming process by which I form that belief is reliable. Then, if warrant is a matter of being produced by a reliable process (as reliabilists tend to argue), I will know G. If in fact the conditions for warrant are more complicated, but still externalist in nature, then I will know G just in case those more complicated conditions are met. In short, on various externalist theories of knowledge, we can explain how it is that we can know causal facts that are neither directly observable themselves nor reducible to facts that are. We can know such facts because our beliefs can satisfy the relevant externalist criteria for warrant, regardless of whether or not those facts or the facts they can be reduced to are present in the content of our perceptual experiences in the way that colors, shapes, and motion are.

The internalist, too, has something to say here.[24] Suppose, again, that a fact G is neither directly observable itself, nor reducible to facts which are all directly observable. It might nevertheless be the case that one can know that G obtains in virtue of an inference to the best

---

[24] I thank Michael Tooley for pointing this out.

explanation. While I'll leave the development of such a suggestion to the internalist, I think we can see that the objector's view of the requirements for causal knowledge is unnecessarily demanding.

IV. Hall's epistemological argument

Ned Hall has recently advanced two arguments against anti-reductionist accounts of causation. One is epistemological and will be treated here; the other will be discussed in section five.

Hall observes that in many cases, indeed perhaps in most, "our knowledge of or justified beliefs about causal relations are inferential."[25] This fact—the fact that we have inferential knowledge (or justified belief) about causal relations—poses a special problem for the anti-reductionist, according to Hall:

> You flip a switch, and a light goes on. You flip the switch back to its original position, and the light goes off. You repeat the experiment a hundred times, and the same correlation obtains. You infer that flipping the switch (say, the first time) *caused* the light to go on. You did not perceive this causal relation. All the same, your inference certainly seems reliable. But it is a mystery why, on the non-reductionist[26] view, it should be.

Let's call the event of your flipping the switch the first time "C", and the event of the light turning on (that same time) "E." Hall's contention is that:

(2) The anti-reductionist view can't explain why the inference from the facts described to the conclusion that C caused E is a reliable inference.

But (2) seems false. Note that as you repeatedly flip the switch, you ascertain that the timing of changes in the position of the switch co-varies precisely with the timing of changes in the illumination state of the light-bulb. (When you flip the switch, the bulb comes on immediately. And when you flip it again, it immediately goes off. And then it doesn't come on again until…you flip the switch, at which point it comes on immediately.) What could be the reason that this co-variation obtains?

---

[25] Ned Hall, "Causation," in *The Oxford Handbook of Contemporary Philosophy*, p. 519. All quotations in this section are from p. 519.
[26] As noted in section two, Hall uses "anti-reductionism" and "non-reductionism" synonymously, at least in this piece.

One logical possibility is this: the co-variation obtains because the bulb is on a timer (so that it blinks on and off), and it just so happened that there was an extraordinary coincidence of timing between your flipping the switch and the light-bulb's own timing mechanism. Needless to say, this is exceedingly unlikely. Another logically coherent but exceedingly unlikely possibility is that someone else is watching you and causing the light-bulb to turn on whenever they see you flip the switch (which is in fact unconnected to the bulb). But there is, of course, another quite obvious possibility: The co-variation obtains because the switch is connected to the bulb in such a way that flipping on the switch will start a causal process that eventuates in the illumination of the bulb and flipping off the switch will block that process. That is, the co-variation obtains because flipping the switch is an indirect cause of a change in the bulb's state.

Anti-reductionism is entirely compatible with this last explanation. On SAR, for example, some events are indirect causes of other events in virtue of being connected by causal chains or processes. And nothing about SAR precludes its being the case that such causal processes are usually started by changes in the states of physical objects. But then an anti-reductionist can explain the reliability of the inference quite easily, as follows: In the situation described, the observed co-variation obtains either because flipping the switch is an indirect cause of a change in the light-bulb's state, or for some other reason. Any other remotely plausible reason (e.g. a coincidence of timing or a hidden prankster) will require the obtaining of a very improbable state of affairs. It is therefore very likely that the co-variation obtains because flipping the switch is an indirect cause of a change in the light-bulb's state. Given that the co-variation does obtain, it follows that it is very likely that flipping the switch is an indirect cause of a change in the light-bulb's state. But then it is very likely that "C caused E," and so the inference to the conclusion that C caused E is reliable.

The existence of this explanation provides adequate reason to reject (2). Although Hall's argument might have force against certain forms of anti-reductionism, it poses no difficulty for SAR.

V. Level-neutrality

Hall points out that an account of causation "needs to be 'level-neutral': it needs to respect

the insight that causal relations obtain at and across a wide-variety of levels of description."[27] He

explains:

> I drink some coffee, and perk up as a result. There is causation at the level of my body: drinking the coffee causes me to perk up. There is causation at the biochemical level, involving the interaction of caffeine molecules with other molecules in my body. There is causation at the atomic and subatomic levels as well.[28]

After making this observation, Hall argues that anti-reductionism (in at least one of its versions)[29]

suffers from a problem relating to the connection between causal relations at different levels:

> [I]t is not that the non-reductionist cannot respect the existence of such relations [causal relations at different levels]; for example, when I drink the coffee there is, on his view, a two-place singular relation of causation that obtains between the event of my drinking and the event of my perking up, and likewise a multitude of instances of this relation at the biochemical level (between events involving caffeine and the cells of my body). What is a complete mystery is why there should be any relations whatsoever of metaphysical dependence between the patterns of instantiation of causation at these different levels. Minimally, it seems that we should have this much: if in any world the pattern of events in my body is exactly as it actually is, and the pattern of instantiation of causal relations at the biochemical level is exactly as it actually is, then so too the events of my drinking the coffee and my perking up instantiate the causal relation. But if all we say about the causal relation is that it is one of the fundamental relations instantiated in our world, then our view provides no resources for explaining why this relation of metaphysical dependence should hold.[30]

The anti-reductionist might indeed have a problem if all he said about "the" causal relation were

that it is one of the fundamental relations instantiated in our world. But he can say more; he can

make the sorts of distinctions made in section two. The anti-reductionist view Hall means to rebut

seems to be a view on which there is just one causal relation, which is ontologically primitive,

and which holds between the drinking of the coffee and the perking up. But from the perspective

of SAR, things should be described differently. It is true that the drinking does cause the perking

up. But what we have here is a relation of indirect production (or some form of indirect causation,

---

[27] Hall, "Causation," p. 515.

[28] Ibid.

[29] This argument is directed at the Tooley-Armstrong brand of anti-reductionism.

[30] "Causation," p. 518.

at least). And (though the concept of derivativeness is conceptually primitive) this causal relation is not fundamental in the sense of being ontologically primitive. Any relation of indirect causation is ontologically reducible to the causal relations between the intermediate causes.

This difference between SAR and the account Hall has in mind provides the resources for explaining the metaphysical dependence between levels of causation. Focus on the relation of metaphysical dependence expressed by the conditional:

(3)   [I]f in any world the pattern of events in my body is exactly as it actually is, and the pattern of instantiation of causal relations at the biochemical level is exactly as it actually is, then so too the events of my drinking the coffee and my perking up stand in the causal relation in which they actually stand.

SAR can explain why this relation of metaphysical dependence holds. The explanation would proceed in two stages. First, letting "C" be the event of drinking the coffee and "E" be the drinker's perking up, C and E would be given more detailed descriptions in chemical and biological terms (e.g., C could be described as the event of ingesting a particular liquid containing significant quantities of caffeine, while some detailed description of the neurophysiological state of becoming more alert would be the more precise description of E.) The second stage of the explanation would involve pointing out that C and E are linked by a chain of indirect causation (ingesting this liquid containing significant quantities of caffeine causes caffeine to be absorbed into the bloodstream, which in turn causes….which in turn brings about the being-perked-up state). Now we can explain why the metaphysical dependence expressed in (3) holds. If the pattern of events in my body and the pattern of instantiation of causal relations at the biochemical level are exactly as they actually are, then it must be that the intermediate causes between C and E are all in place. But then it will follow that C is an indirect cause of E, which is just to say that the events of my drinking the coffee and my perking up stand in the causal relation in which they actually stand.

VI. Conclusion

In this paper I've tried to show that there is a form of anti-reductionism which is not threatened by the unobservability argument or the two problems raised by Hall. I want to close with a brief comment about the implications of anti-reductionism for future work on causation. There has been a great deal of writing on the relation between causation and probability-increase, and on quite complex methods for identifying causes via statistical methods. An anti-reductionist can be enthusiastic about these sorts of projects, and about projects mapping out the relations between causation and the other conceptual connotations of causation—counterfactuals, explanation, etc. Such projects can give us greater clarity about the conceptual relations between causation and these closely connected concepts, and perhaps greater clarity about the different types of causation. So an endorsement of anti-reductionism need not be "defeatist," in that it need not involve a negative attitude towards further work on causation.[31]

Michael Rota

University of St. Thomas