ANJAN CHAKRAVARTTY

CAUSAL REALISM: EVENTS AND PROCESSES

ABSTRACT. Minimally, causal realism (as understood here) is the view that accounts of causation in terms of mere, regular or probabilistic conjunction are unsatisfactory, and that causal phenomena are correctly associated with some form of de re necessity. Classic arguments, however, some of which date back to Sextus Empiricus and have appeared many times since, including famously in Russell, suggest that the very notion of causal realism is incoherent. In this paper I argue that if such objections seem compelling, it is only because everyday expressions concerning causal phenomena are misleading with respect to certain metaphysical details. These expressions generally make reference to the relations of events or states of affairs, but ignore or obscure the role played by causal properties. I argue that on a proposed alternative, an analysis in terms of causal processes, more refined descriptions of causal phenomena escape the charge of incoherence. Causal necessity is here located in the relations of causal properties. I distinguish this view from the recent process theories of Salmon and Dowe, which are disinterested in causal realism.

1. CAUSAL REALISM

Theorists about causation agree on precious little, but all are amused by Russell’s inflammatory introduction to ‘On the Notion of Cause’, in which he claims that the idea of causation, ‘like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm’. This shared comic relief, however, masks a common misrepresentation of Russell’s position. It is often suggested that the primary motivation behind Russell’s attack is the idea that the notion of causation is of no use to the sciences, but this is plainly not what Russell believed. Elsewhere he uses causal language freely, and was at pains in later work (e.g. 1948) to give an account of causal phenomena in terms of what he called ‘causal lines’. This was no about-face; these different parts of the Russellian corpus are consistent with one another. His target in ‘On the Notion of Cause’ is not causation simpliciter, but rather ‘the law of causality’ – the philosopher’s notion of causation. It is
this putatively philosophical creation for which he thought the sciences have no use. One aspect of this creation, for lack of a better term, is what I will refer to as ‘causal realism’.

Realism about causation requires two things. First, according to the realist, causation is objective, meaning that it is something that occurs in an “external reality” as opposed to something that is merely subjective, a feature of our thoughts or perceptions alone (that is, merely an idea or a concept). The distinction between objective and subjective causation thus concerns the issue of whether or not causation is mind-independent. Second, according to the realist, causation involves some sort of necessity with respect to the connection between causes and effects. This element of realism is neatly captured by Russell’s phrase ‘law of causality’. Though many empiricists dispute that it should, the idea of a law often connotes some form of necessity. What manner of necessity is a question answered in different ways by different causal realists, but all sign up to it in some form or other. For now, let it suffice to say that by invoking necessity, realists of different stripes maintain that there is more to causation than mere constant or probabilistic conjunctions of events. Merely subjective accounts of causation hold that if there is such a thing as causal necessity, it is an idea or a concept only. Objective accounts hold that there is such a thing, and that it is a feature of the world quite apart from our ideas or concepts. In Section 4, I will make some more specific remarks about proposals for causal necessity.

The two most celebrated rejections of causal realism are familiarly attributed to Hume and Kant respectively. One might, in the tradition associated with Hume, reject anything more metaphysically weighty than simple succession, and explain away the notion of causal necessity as a psychological figment projected without warrant onto nature. Hume famously denied one route to a belief in objective causal necessity, the route of perception, and he was of course correct; we do not have impressions of necessary connections. Admittedly, some argue that we have some such impressions in experiences of bodily forces. Fales (1990, pp. 11–25), for example, notes that when we engage in activities involving pushing or pulling, or being pushed or pulled, we experience sensations of force or power, and Hume (1975/1777, p. 67) makes similar observations regarding impressions of bodily force or exertion (‘nisus’). But many putatively causal situations offer no such impressions and more importantly, impressions of force are irrelevant here, for they are not tantamount to impressions of necessary connections. Forces and causal necessity are different things, and only the former are amenable to sensation.
Causal necessity is unobservable, and unlike many scientific unobservables, it is not even a possible object of detection.

Though they need not be, scientific realists are often causal realists as well, and this I suspect is no accident. The project of scientific realism is partially constituted by attempts to furnish explanations for observable phenomena by positing things underlying them, generally by inference to the best explanation. If such inferences lead to a knowledge of the existence of electrons and genes, they might also lead to a knowledge of the existence of causal necessity. This is merely suggestive, of course, since there are disanalogies between scientific unobservables and the posits of metaphysics, but the suggestion remains. The mere existence of regular patterns of behaviour may be sufficient for purposes of prediction, but not for the purpose of explaining why there are such patterns to begin with. A notion of objective causal necessity allows the realist to distinguish between causal regularities and merely accidental ones.

A similar appeal to scientific realist sensibilities illuminates the causal realist response to Kantian scepticism. One might, in the tradition of Kant, hold that causation is part of a basic conceptual apparatus possessed a priori, required for the very possibility of experience. Do not ask, however, whether these concepts apply to the noumena, which are beyond that which we make intelligible. For Kant, the concept of causality is required for the judgment that an event has taken place. It is by applying the ‘rule of cause and effect’ that we experience causal phenomena. Causal necessity is thus housed in a category of the understanding. It is a purely representational aspect of things, not a feature of things-in-themselves. The mysterious dependence of the phenomena (knowable) on the noumena (unknowable) is problematic for scientific realists. Since the dependence is not causal, the noumena are robbed of the explanatory virtue they might otherwise possess. Lacking a clear explanatory role, things-in-themselves are susceptible to scepticism. With no grip on the noumena, the realist sees transcendental idealism collapse into idealism simpliciter.

The Humean and Kantian objections to causal realism have been widely discussed, and it is not my intention to add to those discussions here. Even if a causal realist is satisfied with her reasons for resisting forms of empiricism and transcendental idealism, a third class of objections remains, and my present goal is to focus on these more specifically. This third class of arguments does not take issue with the realist by adopting what might be regarded as incompatible positions ex ante, but rather by attacking the causal realist on her
own terms. One might doubt, starting from the assertion of causal realism, that the position is coherent by its own lights. Perhaps having an incoherent idea or even an incoherent category of the understanding might not be intolerable. We are, after all, less than perfect. But if causal realism requires that we attribute an incoherent phenomenon to a mind-independent world, clearly causal realists are in trouble. It is precisely this sort of worry that was pressed by the ancient sceptics, and that Russell exploits in his attack on the philosopher’s ‘law of causality’. It is this worry that I will attempt to lay out, consider, and dissolve here.

Arguments for the incoherence of causal realism have a venerable history in western philosophy, going back at least as far as Sextus Empiricus (Mates (1996, pp. 175–177)). Russell’s (1953/1918) formulation is something of a 20th-century locus classicus, and his own later account of causation significantly inspires the more recent views of Salmon, Dowe, and others which I consider in Section 5. The Russellian critique has two primary targets: the idea of causal necessity (1953/1918, pp. 174, 177, 183), and the Humean definition of causation in terms of constant conjunction (pp. 174–175). His main arguments for incoherence arise in the latter context but quickly become relevant to the former (pp. 177, 183), which is my main concern here. In the next section I lay out generalized versions of the arguments for the incoherence of causal realism. In Section 3, I argue that the standard realist account, according to which causal relations obtain between events, cannot disarm the incoherence arguments. In Section 4, I propose a different analysis of causation, a “process” theory, and suggest that not only is it immune to worries about incoherence, but that it facilitates various answers to the question of how the realist should understand de re necessity in the context of causation. Finally, in Section 5, I briefly consider other currently available process theories in order to clarify the extent to which they have a different motivation, and thus focus on different questions, than the account given here.

2. ARGUMENTS FOR THE INCOHERENCE OF CAUSAL REALISM

Consider the putative circumstance of one event, $A$, causing another, $B$. The potential problems for causal realism here are threefold. I will refer to these challenges as the ‘contiguity objection’, the ‘regress objection’, and ‘the demand for a causal mechanism’. In each case the
assumption of \textit{de re} necessity will turn out to be central in drawing out apparently problematic consequences.

2.1. \textit{The Contiguity Objection}

In order that $A$ cause $B$, $A$ and $B$ must be contiguous in time. (A variation of this argument might begin with the supposition that causally related events must be spatially contiguous as well; leaving aside issues concerning the spatial demarcation of events, I here focus on the temporal case). Granted, we often refer to events as causally related even when they are not temporally contiguous. Knocking over the first domino in a sequence is a cause of the falling over of the last one, for example, despite the fact that the events of the first and last fall are not contiguous in time. But in these cases we say that there is a causal chain of events connecting the events we originally cite – we appeal to intermediate causes and effects – and hold that each link in the chain is temporally contiguous with the next. So for $A$ to bring about $B$ causally, not mediated by other events, but directly, $A$ and $B$ must be contiguous in time. But $A$ and $B$ cannot be contiguous, because time is dense. That is, between any two instants, say that at which $A$ terminates and that at which $B$ begins, we can always find further instants. Therefore, it is impossible for successive events to be temporally contiguous. Thus, $A$ cannot cause $B$.

Note that the reductio offered by the contiguity objection invokes two explicit premises. The first is the premise that causally related events must be contiguous, and the second is the premise that time is dense. Both of these premises seem well-founded, and at the very least hard to resist. Though some have speculated that time may be discrete or quantized (say, at the Planck scale), the idea that time is continuous and thus dense is an important conceptual resource, allowing us to solve otherwise troubling puzzles such as Zeno’s paradox of the arrow, and to define instantaneously possessed properties such as velocities and accelerations. Topological studies of dimensionality yield the result that discrete time has dimension zero, which conflicts with the standard assumption that time is one-dimensional.\textsuperscript{5} And what of the premise of contiguity? This may be particularly difficult for the causal realist to deny. If events are the relata of causation, it is difficult to see how their connection could be necessary were they not contiguous. If a putative cause $A$ is not, indeed, contiguous with its effect $B$, it is possible that something might intervene subsequent to $A$ so as to
prevent $B$. This is why the objection concerns direct as opposed to mediated causation, for it is generally assumed that given a causal chain of events, relations of necessity obtain only between contiguous events, not between events separated by others in a causal chain. Necessity here means sufficiency; I will return to this point momentarily.

2.2. The Regress Objection

A second potential worry adopts a tactic similar to the first, but this time, rather than focusing on the connection between the putative cause and effect events, $A$ and $B$, the objection targets connections between the temporal parts of $A$ and $B$ themselves. As events, $A$ and $B$ are comprised of changes in an object or objects, but if they are changes there is, presumably, a causal relation between their earlier and later parts, and here we have the makings of a regress. (Let us grant this conception of events as changes for the sake of argument, for even if we allow that events might include unchanging exemplifications of properties at times, many if not most of what are generally regarded as causes and effects are indeed changes.) If an earlier part of $A$ is the cause of the part remaining, then the earlier part cannot serve as the proximate (direct, unmediated) cause of $B$; only the later part of $A$ can serve this role. But whatever duration of $A$ we put forward as containing the proximate cause of $B$, we may then ask about the causal relation between the earlier and later parts of the occurrent change in this duration. In the hope of arriving at a proximate event to serve as a cause, we may diminish the originally supposed cause $A$ without limit. The same sort of difficulty plagues the identification of the proximate effect, $B$, of $A$.

In order to escape these difficulties, one might deny the premise that time is infinitely divisible, but infinite divisibility is entailed by continuous time, and we have already noted that there are reasons for assuming a continuum. And once again it seems that the causal realist’s commitment to de re necessity is what fuels the objection, for if one does not view the connection between causes and effects as necessary, there is no obvious need for a workable understanding of proximate causation (though one might desire such a thing for independent reasons). In order to see why this is so, it is important to appreciate, I think, an interesting difference in the conceptions of necessity at work in the contiguity and regress objections, so let us consider this difference briefly.
The contiguity objection derives its force from the causal realist’s assumption of necessity understood in terms of sufficiency. The worry is that if A and B are not contiguous, there can be no guarantee that A is causally sufficient, given the possibility of intervening events preventing B; necessity is lost, for B might not occur following A. The regress objection, on the other hand, preys on an assumption of necessity that is best understood not in terms of sufficiency, but rather in terms of necessary conditions. Lewis (1973) has this sort of assumption in mind when he argues that ‘causal dependence’ – the idea that if A did not happen, B would not happen either – is not transitive. That is, says Lewis, the truth of the counterfactual conditional (‘if A did not happen, B would not happen either’) indicates a causal dependence only in cases of contiguous events. If A causes B, and B causes C, A is a mediated cause of C, but it would be a mistake to say that generally, C is caused dependent on A, since there are causal scenarios in which A is not a necessary condition for C; B might have come about some other way. A similar point can be made regarding effects: presumably it is often the case that causes furnish necessary conditions for their proximate effects only, not the temporally more distant parts of effects. Lewis’ analysis of causation is rightly controversial, but recalling it is helpful in illustrating the intuition targeted by the regress objection. It is precisely the hope of finding unmediated causes and effects, to which one might think the necessity of causal dependence applies, that leaves the realist susceptible to a regress which attempts to dismiss all parts of A and B save those serving as proximate cause and effect.

Here is another illustration, this time by way of an analogy. J. L. Mackie (1965) famously analyzes causation in terms of what he calls ‘INUS’ conditions. All of the conditions that are jointly sufficient to bring about an effect, B, comprise a complex condition, which is itself generally unnecessary, given that B might have come about some other way. But when we cite a cause, we are primarily interested in picking out some component part of this complex: viz., a part that is necessary in bringing about B, but that is insufficient on its own. In other words, causes are Insufficient but Necessary parts of Sufficient but Unnecessary complex conditions. Consider Mackie’s ‘N’; when we identify a cause, we are picking out some part of the conditions preceding B that is necessary. It is an analogous intuition on the part of the causal realist that is exploited by the regress objection. The realist is invited to associate causal necessity with that part of A that is necessary to bring about B. And in searching for that part, she is carried along on a hopeless quest for proximate causes and effects.
2.3. The Demand for a Causal Mechanism

The contiguity and regress objections both exploit the idea that time is continuous, thus entailing that it is dense and infinitely divisible. A third objection, however, constitutes a more general concern, and does not depend on any assumptions with respect to the nature of time. The concern is general, but as a specific means of introduction, let us consider a scenario that may escape the regress objection. Imagine a case in which we are able to diminish the event that is to serve as the proximate cause, $A$, to such an extent that it contains no change between its earlier and later parts. In other words, let us consider not a change at all but rather some static, unchanging state of affairs preceding $B$. One might then ask, if $A$ is no longer any sort of change, how does it give rise to an effect? How is it that something static should suddenly bring about a change, when it itself has no element of change within it? How can something that is not a change bring about anything? Questions such as these are symptomatic of a more general worry concerning what is happening, precisely, when one thing is thought to give rise to another causally. On further reflection, it appears that this concern is not limited to cases in which effects are imagined to follow from static states of affairs, but applies to any causal succession. The worrier is seeking some sort of mechanism for the connection between causes and effects on which to hang the idea of causal necessity.

The challenge to describe the nature of this connection has, historically, resisted any detailed response. Metaphors abound: links, chains, ties, glue, cement, bringing things about, and perhaps most highly scorned, the “powers” of ancient metaphysics which “give rise” to subsequent phenomena. The notion of a causal connection is barely intelligible, it is claimed; if there is something to objective causation over and above mere regularity, we have yet to furnish anything like a helpful, qualitative description of it. The complaint that causal realism is empty if the realist’s analysis of causation is incapable of describing the nature of a sui generis mechanism in which causal necessity inheres is part of the traditional empiricist critique. For “nothing will do just as well as something about which nothing can be said”.7 We have no decent conception of what a causal connection is, so if this is what the idea of objective causal necessity amounts to, ultimately, we gain nothing worth having by postulating it.

The three objections outlined above offer challenges to the coherence of causal realism on its own terms. That is, they are not
objections based on Humean scepticism regarding things not ame-
nable to sensation, nor Kantian scepticism based on the status of
causation as purely a category of the understanding. Rather, they
assume the standard realist picture of causation as a mind-indepen-
dent relation between events, imbued with some form of necessity,
and offer reductio ad absurdum arguments in the case of the contiguity
objection and the regress objection, and an unanswered challenge in
the case of the demand for a causal mechanism. On the account of
causation that these objections assume, their challenges are not easily
met. In Section 4, however, I suggest an analysis of causation on
which the objections of contiguity and regress do not arise. This
account also partially addresses the demand for a causal mechanism.
I suspect that a complete response is beyond the reach of any account
of causation, in principle, but that this offers no genuine impediment
to causal realism. First, however, let me make the case that standard
realism cannot respond to the charge of incoherence.

3. WHY EVENTS-BASED ACCOUNTS CANNOT ANSWER THE CHARGE

If incoherence-type arguments seem worrying for causal
realism – particularly the objections of contiguity and regress – it is
only, I think, because everyday descriptions of causal phenomena are
ambiguous with respect to the precise details of causation. This
ambiguity finds a home in the common practice of identifying events
as the principal actors in causal relations. Not all realists, however,
would agree. The standard response to incoherence objections has
been to claim that they can in fact be answered on an events-based
account of causation. In this section, I consider the three responses
commonly suggested by causal realists. I will argue that the only
reasonable interpretations of the first are that either it offers no
response at all, or that it can be charitably reconstructed along the
lines of the second. The second and third responses are suggestive,
but on further consideration we will find that in fact their promise
leads realists to the proposal of Section 4.

Consider the contiguity objection. The first strategy for dealing
with the charge of incoherence on an events-based account is to
contest the notion that there is anything problematic in saying that
a putative cause event, \( A \), and its putative effect, \( B \), are contiguous.
Beauchamp and Rosenberg (1981), for example, say this in a dis-
cussion of Russell’s formulation of the problem (in defence of their
Humean account of causation). Contra Russell, they argue, two
events can be ‘both contiguous and successive if the first begins at instant \( t_1 \) and ends at instant \( t_2 \), while the second begins at \( t_2 \)’ (p. 196). Many realists, I suspect, implicitly take something like this contention for granted. But more careful consideration reveals that this cannot stand as a convincing response without further elaboration. To say merely that \( t_2 \) marks a point in time at which \( A \) ends and \( B \) begins is misleading, because it trades on the ambiguity of beginnings and endings.

This ambiguity is illuminated by thinking more carefully about the contiguity objection. The argument here assumes that events are discrete, meaning that they can be defined on closed temporal intervals. Consider a series of temporally ordered instants: \( t_1, t_2, t_3, \ldots \). To say that a putative cause event \( A \), for example, is defined on the closed interval \([t_1, t_2]\), is to say that \( A \) contains (includes) both instants \( t_1 \) and \( t_2 \), but none before \( t_1 \) or after \( t_2 \). Thus it is immediately apparent why the first response is insufficient as it stands, for if both \( A \) and \( B \) contain \( t_2 \), the events overlap, in which case they are not strictly successive. If only one or neither of \( A \) or \( B \) contains \( t_2 \), the events cannot be contiguous, given that events are discrete and that time is dense, for there will always be instants between them. What, then, is intended by the realist claim that \( A \) ends at \( t_2 \), and that this is when \( B \) begins? The two interpretations outlined in this paragraph appear to be exhaustive, which would indicate that this first answer to the contiguity objection offers no response at all.

Perhaps this is too quick — perhaps there is something else the realist might intend, here. While the two scenarios described above are indeed exhaustive given the assumptions of the contiguity objection, the realist might be able to reformulate the first response by rejecting one or more of these assumptions. More specifically, the premise that events are discrete entities invites further scrutiny, and it is precisely this line of inquiry that motivates the second realist response. Imagine a putative cause, \( A \), not on the model of causal relata presupposed by the contiguity objection, but rather on a continuous model. That is, let us define \( A \), not as a discrete entity, but as a continuous entity. For example, rather than defining \( A \) on the closed interval \([t_1, t_2]\), we might define it on the half-open interval \([t_1, t_2)\); \( A \) would then contain all instants from and including \( t_1 \), up to but not including \( t_2 \). \( B \), the putative effect of \( A \), could then be defined on the interval \([t_2, t_3]\), and so on.\(^8\) In this way, causerally related events could be thought of as successive and contiguous, for there are no instants in time between \( A \) and \( B \).

I believe that this second realist response is on the right track. The key to answering incoherence-type objections rests in conceiving of
causation, not as a relation between discrete entities, but by in some way appealing to properties of the continuum. The specific idea of conceiving of events in a continuous manner, however does not go quite far enough. There are, I suggest, at least two difficulties with this proposal that push us in the direction of a different account of causation. While neither of these difficulties represents a knock-down argument against the standard realist picture, they do undermine the idea that it is relations between events on which the realist should hang her commitment to objective causal necessity. Let us consider these points in turn.

The first problem with the realist’s second response is that, while defining events on continuous intervals may counter the contiguity objection, it offers no help with the regress objection. The latter, recall, argues the impossibility of there being proximate causes and effects, given that any such candidates may be diminished ad infinitum. Some, like Mellor (1995), may be willing to accept this consequence – perhaps causation, like time, is dense, and there is no such thing as proximate causation. If the realist goes this route, however, she may have to give up causal necessity in the form of necessary conditions as part of her general account of causal realism; as discussed in Section 2, this form of necessity does not apply generally to non-proximate causation. Furthermore, if causation is dense, it is arguable that part of the motivation for understanding causation as a relation between events is lost, for then there is no sense in talking about events directly causing or being caused by others. That is to say, the idea that events qua ontological category of entity are the fundamental relata of the causal relation is undermined, for now there is nothing like unmediated causation between events. No event $A$ can serve as a direct cause of another, $B$.

This suggests another, closely related and I think more serious problem with the realist strategy of defining events on continuous intervals. The upshot of this problem for the realist is not to reject the strategy outright, but simply to downplay the emphasis on events in a realist analysis of causation. Once we see that on the realist’s second response to incoherence-type arguments, there are no such things as proximate causes and effects, the demarcation of events to serve as the relata of causation becomes a fairly arbitrary matter. We are at liberty to define $A$ on $[t_1, t_2)$ and $B$ on $[t_2, t_3)$ if we wish, but we could just as easily choose differently without incorrectly describing the causal facts of the matter. That is, there is no constraint here based on correctly describing some mind-independent entities, $A$, $B$, and their causal relation. We might just as well choose to identify $A$ on a
slightly different half-open interval, and do likewise with $B$. Any choice will do so long as $A$ is defined on an interval that is open up to the instant at which $B$ definitively begins. Here again, the idea that there is a special ontological category of entities, events, that stand in privileged, causal relations, should begin to ring hollow in realist ears. For these events can be sliced up any way we please. What is crucial here is the continuum along which causation occurs, not any particular temporal slices we may, for whatever pragmatic reasons, choose to recognize as events. The conventionality of the choice seems to render our explanatory reliance on events a pragmatic feature of how human beings, for explanatory purposes, decide to divide up the continuum of happenings. But this is some distance from the idea that particular relations between particular events constitute an objective, mind-independent thing called causation.

Finally, let us consider a third and last response to the charge of incoherence on the part of those whose causal realism focuses on relations between events. Reflect for a moment once again on the contiguity objection. Some will be tempted to complain that the objection simply misses the point: it is irrelevant whether $A$ and $B$ are strictly contiguous; the question of whether something could intervene between events $A$ and $B$ so as to prevent $B$ from coming about is a red herring. In the circumstances, there is no intervention. In the circumstances, $A$ is sufficient and/or necessary (depending on how $A$ is characterized, precisely) to bring about $B$. In other words, in any given case, by ‘$A$’ we mean to refer to a collection of factors that is relevant to bringing about $B$, and which, as it happens, excludes factors that would preclude $B$.

Two things should be noted here. First, by appealing to what is causally necessary ‘in the circumstances’, this response gives an analysis of causation that is wholly singular. This observation also applies to the second realist answer to the charge of incoherence, in which events are defined on continuous temporal intervals. Here too necessity will be explained in terms of the specific circumstances that obtain. These answers focus on the conditions preceding $B$ in concrete cases, but in doing so describe very particular sets of circumstances. Considering the weight given here to the presence of specific conditions and the absence of intervening ones, it is unclear how these views might be generalized so as to yield an account of general causation; that is, an account in terms of laws. For it is unclear what the identity criteria should be for the types of events that could stand in such laws.

A second point about this last events-based response to the contiguity objection is closely related to the first. What this response
does, in fact, is to relocate the idea of causal necessity from the sphere of relations between events *per se*, to that of various circumstances that in part compose particular events. It effectively downplays the traditional realist’s emphasis on events, and now turns a spotlight on specific conditions that obtain or are absent. In other words, in order to give promising responses to charges of incoherence, the realist has shifted from taking about relations between events to talking about specific combinations of *properties*. This insight forms the basis of the proposal in Section 4. And as we shall see, by changing the emphasis from relations between events to a finer-grained analysis of relations between properties, various options for an account of general causation emerge immediately from the context of singular causation.

4. CAUSAL PROCESSES AND THE COHERENCE OF CAUSAL REALISM

Earlier I suggested that incoherence-type arguments may seem compelling only because of an ambiguity in traditional realist characterizations of causation as a relation between events. Though misleading, however, the traditional realist picture is not in the wrong ballpark entirely. The problem with the traditional picture is that it privileges the role of events in giving an analysis of causation, and this pays insufficient attention to the precise metaphysical details. Focusing on events has the unfortunate consequence of obscuring the role played by those properties of things that we take to explain their causal behaviours. I will call these properties ‘causal properties’ to distinguish them from other possible non-causal ones such as mathematical, logical, and epiphenomenal properties. Once we appreciate the role of causal properties, I will argue, the kinds of worries that give rise to incoherence objections simply do not arise in the first place. This understanding provides a framework for causal realism, and vindicates it against the charge of incoherence.

When I claim that incoherence objections are premised on a realist account that pays insufficient attention to certain details of causation, I have in mind specifically the objections of contiguity and regress. In spelling out the details and rising above these objections, however, we will also find a partial response to the demand for a causal mechanism. Let us turn to these details now. Descriptions of causal phenomena in terms of relations between events are useful for many purposes, but it is not events *qua* events that do the work of causation. Events participate in causation *because* they incorporate causal properties of objects. Referring to
events as the relata of causation makes sense of much of our phenomenal experience simply because, as it happens, these things harbour the ontological ingredients, causal properties, that are ultimately responsible for causal phenomena. When we say that event \( A \) causes event \( B \), what we are doing is employing a coarse-grained shorthand for the details of causal interaction. This shorthand works rather well for most everyday and scientific purposes. It manages to latch on to the details in such a way as to do justice to our coarse-grained observations of causal activity.

So what does it mean to say that causal properties ‘do the work’ of causation? In order to answer this question, let us consider in some detail what these sorts of properties are, exactly. A causal property confers dispositions for behaviour. That is, a causal property is one that confers dispositions on the objects that have them to behave in certain ways when in the presence or absence of other objects with causal properties of their own. The property of mass, for example, confers on bodies that have it certain dispositions to be accelerated under applied forces. The property of volume on the part of a gas confers certain dispositions to vary in temperature in ways correlated with applied pressures. Causal phenomena are produced by the ways in which objects with properties are disposed to act in consort with others, and it is this fact that realists should exploit to answer the charge that causal realism is incoherent.

Recall that in response to the objections of contiguity and regress, the events-based realist account made two promising but self-undermining moves. The first was to appeal to properties of the continuum in order to avoid difficulties engendered by the temporal relations of discrete events. The second was to shift from talking about events, \( \text{per se} \), to talking about collections of causal properties. Let us now follow these moves and see where they lead. Here is a first pass: objects with causal properties are engaged in continuous processes of interaction. Dispositions borne by objects in virtue of their properties are continuously manifested in accordance with the presence and absence of other objects and properties. Objects with causal properties are thus in a continuous state of causal interaction, a state in which relations between causal properties obtain. For example, a volume of gas that comes into contact with a source of heat may expand in virtue of the dispositions afforded such volumes by properties such as temperature and pressure, and in doing so will come into contact with other regions of space. The properties present in these new regions together with those of the gas will determine how both are further affected, and so on.
Now if this sort of process is what causation is all about, it seems only reasonable that we be able to explain why realists generally take events to be the relata of causation. But we have stumbled across this explanation already. Of the continuous flux of causal activity that surrounds us, we take notice of only certain parts: ones that interest us, or that are useful in the context of pursuing particular objectives, such as realizing desired states of affairs or avoiding harmful ones. And generally speaking, we describe these parts in terms that are consistent with the coarseness of our sensory appraisals of the relevant phenomena. Thus, causation is often described crudely in terms of events. But we should be wary of fixing ontological commitment simply on the basis of the grammatical form of these descriptions. Talk of events as the relata of the causal relation – ‘A causes B’ – is elliptical for descriptions of aspects of continuous processes: ‘A precedes B, and the object(s) involved in A have dispositions, some of whose manifestations are present in B’. Events that are changes involve changes in properties, certainly, but any given event generally overlaps multitudes of changes in the properties of the objects concerned. Events occupy time slices during which objects with causal properties are engaged in continuous processes of causal interaction.

Perhaps another example will help to clarify the notion of an understanding of causation in terms of continuous processes. Take the classic example of a collision of two billiard balls. The first ball moves toward the second, which is at rest. As the balls collide, the first stops, and the second moves on. Here we have event A, the motion of the first ball, and event B, the motion of the second. A is thought to cause B. But here the objections of contiguity and regress apply. According to the contiguity objection, A and B cannot be temporally contiguous. For given that the motion of the first ball has an ending and the motion of the second has a beginning, and that the former precedes the latter, there are instants in time between A and B. According to the regress objection, the earlier motion of the first ball cannot be the proximate cause of the motion of the second; we may consider the motion of the first ball in shorter and shorter durations prior to the termination of A without limit in search of a direct cause. Similarly, we might truncate B infinitely in search of a proximate effect.

Now let us increase the resolution of the analysis. What the contiguity objection takes to be a discrete event, the motion of the first ball, actually overlaps a continuous evolution in various properties of the ball, as some of the many dispositions it has are manifested in light of other properties with which the ball “comes into
contact” – those of the table, surrounding air molecules, and so on. As the first ball approaches the second, they interact in virtue of the properties they each possess: velocities, momenta, etc. These interactions continue in the form of further continuous evolution in the relevant properties (and thus continuous alterations in the relevant motions) until such time as the balls are no longer within a causally efficacious range of one another. Series of discrete events are here replaced in the analysis by a continuous alteration of properties, each conferring dispositions for behaviour on the objects possessing them. Thus, worries about temporal contiguity between discrete, successive events are replaced by an acknowledgement of continuous processes of causal interaction. The search for events to serve as proximate cause and effect is replaced by the understanding that candidates for these simply constitute convenient or conventional divisions of the continuum of happenings into otherwise arbitrary time slices, themselves inhabited by numerous causal interactions. A more finely-tuned analysis of causation leaves aside the objections of contiguity and regress.

Furthermore, this process view is flexible enough to provide a framework for the analysis of causation not merely in cases of causal change, but also in cases involving static states of affairs that some think to be causal as well. Here the analysis would be given in terms of equilibrium relations, manifested in accordance with dispositions conferred by the relevant properties. Indeed, one might even extend this to an analysis of the controversial case of simultaneous causation, where one state or event appears to cause another, co-temporal one, or even cases in which two or more co-temporal aspects of the same state cause one another. I will not offer an argument for simultaneous causation here; in fact, I suspect that such putative cases are always describable in non-simultaneous terms. Let me sketch nonetheless how the proposed causal realism would explicate simultaneous causation for those inclined to do so. Imagine a case in which one might say that what causes a gas to occupy a particular, constant volume is that it has, co-temporally, a particular temperature and pressure. One could just as well say that what causes the gas to be at a constant temperature (pressure) is that it has, co-temporally, a particular volume and pressure (temperature). Alternatively, one might explain why the gas has a certain temperature (pressure) by appealing to the dispositions conferred by its prior volume and pressure (temperature). Either way, various causal properties of the gas are stable, in keeping with the dispositions they confer on the gas itself.
Still, worries about the contiguity and regress objections may persist. One might wonder what prevents the sceptic from urging these same arguments once again, this time with continuous processes of interaction as their target. But recall that the contiguity objection exploits the premise that causal relata are discrete. By adopting the idea that causation is a process, the realist is trading in an analysis in terms of successions of discrete events for one in terms of continuous processes. “Trading in” is perhaps a more radical suggestion than one might think. Fully to embrace causation as a continuous process is to view processes as causally fundamental or basic. Making the switch is not a matter of simply replacing large-scale events in the analysis of causation with micro-events made up of changes in the properties of objects. For the realist to think of causation as a process is to preclude discussion of causation as comprised of relations between discrete events, except as elliptical for descriptions of aspects of processes. That is why the contiguity objection does not arise here. Of course, this does not prevent us from carving events out of parts of causal processes and calling them causes and effects for the sake of convenience. We do this very effectively, as our success in facilitating everyday and scientific tasks confirms. But realists should understand that when we speak of causal relations between events, their causal status derives from the fact that they overlap continuous interactions that together comprise processes.

On the process view, the causal realist is thus immune to the objections of contiguity and regress. What about the demand for a causal mechanism? The sceptic here is asking for a deeper explanation of what precisely is taking place when one thing is thought to give rise to another causally. This deeper explanation might include, for example, a description of some sort of mechanism for the link between causes and effects. This demand for a causal mechanism, however, is only partially addressed by the metaphysics of causal properties. The most one can say, I think, is that causally efficacious events incorporate objects with property-conferred dispositions, and the occurrence of subsequent effects can thus be understood in terms of manifestations of the relevant dispositions of the objects involved. The process account, while de-emphasizing the role of successive events in making sense of objective causal necessity, should not, however, be construed as uninterested in the fact that some alterations follow others in time. On the contrary, what we are often most interested in are the ways in which the states of objects evolve. States change; on the proposed view of causation, however, explaining
precisely how such change occurs is something we can say only so much about.

This is to concede Hume’s point that ultimately, we have nothing like a “picture” of what is happening when one thing brings about another, beyond that which is observable. That is why the demand for a causal mechanism cannot be fully satisfied. Objects with causal properties are disposed to behave in certain ways when in the presence or absence of other objects with properties of their own. Causal phenomena are produced by the ways in which property-conferring dispositions are linked to one another, and noting this may be the best that we can do. As far as the realist is concerned, however, this best is good enough, and to think that the sceptic’s demand for a causal mechanism is problematic for realism simply begs the question. The empiricist, for example, advocates a rejection of causal realism in part because of the suggested obscurity of causal necessity, and appealing to dispositions, she will claim, does little to clarify the situation. But exercising these sorts of epistemic principles is simply part of what it is to be an empiricist. Conversely, realists generally favour different epistemic principles – ones that, as I hinted in Section 1, do not require that the demand for a causal mechanism be satisfied.

As in Section 1, let us first consider an inferential practice that is characteristic of scientific realism. Although different in important respects, the scientific case helpfully illuminates the sort of inference at work in the metaphysical case. Many of the unobservables whose existence scientific realists routinely infer on the basis of their explanatory virtues are grasped at best metaphorically. When we ask for a description of an electron, not in terms of what we interpret as empirical measures of its quantifiable properties, but rather in terms of a deeper, qualitative picture, we are told to think of it loosely in terms of several (and jointly inconsistent) models: particles, waves, clouds, etc. The metaphors in terms of which we fill out our conceptual picture of how things give rise to other things causally are also vague, and this is an interesting epistemic fact, but not one that, for the realist, has any definite ontological implications. Realists do not exclude unobservables that play important explanatory roles on the grounds that the metaphors they employ to conceptualize them are less precise, descriptively, than empiricists would otherwise like.

An electron is no ordinary particle, wave, or cloud. It is an entity whose properties stand in certain relations, and it is our grasp of the relations between these and other causal properties that allows us to
explain the electron-related phenomena we detect and create in a scientific context. An analogous point can be made by causal realists regarding necessity. Here too we employ metaphors. What is the notion of a power, after all, if not an attempt to render intelligible, metaphorically, the notion of causal necessitation? The explanatory benefit of causal realism is crucial, argues the realist, in meeting the demand for an explanation of why non-accidental regularities occur. Those tempted to complain that we have no precise “picture” of causal “bringing about” are looking for precision in the wrong place. Of course, there is an important sense in which the case of scientific unobservables and that of causal necessity differ. It would be peculiar – a category mistake – to think of causal necessity as an entity, as we do an electron. Despite this, realists maintain that causal necessity is de re, a feature of processes in the world, not merely our descriptions of them. How, then, is this de re character to be understood?

One much discussed proposal for understanding necessity here stems from the idea that causal laws (and laws in general) are relations between properties. This approach is commonly associated with Dretske, Tooley, and Armstrong. Although they vary in ontological commitment and other details, each of these authors gives an account of the necessity that characterizes certain relations between properties. Another approach is taken by Shoemaker, who argues that what makes a causal property the property that it is, are the dispositions for behaviour it confers on the objects that have it. Causal necessity then follows from the nature of property identity: objects that behave differently in exactly similar circumstances in any possible world have, ipso facto, different causal properties. On both of these approaches, instances of singular causation are closely linked to general or law-like causation, since properties and relations instantiated in a given instance may also be present at other times and places and involving different particulars, thus generating regularities.

All of these views are, indeed, controversial. But it was never my intention to settle these further questions here. We began with the claim that the very idea of causal realism is incoherent, and what I hope to have shown is that by understanding causation in terms of continuous processes, the realist is safe from the incoherence arguments that plague events-based accounts of causation. By relocating necessity from relations between discrete, successive entities like events, to the relations of causal properties, the realist about causation takes one step forward. The question of what further position
might best describe objective causal necessity, whether in ways suggested by the proposals mentioned above or otherwise, must wait for another occasion. Sceptics about causal realism have things to say about the plausibility, and perhaps even the coherence of these views as well. I hope, however, to have shown that this further debate can and should take place, free of some of the classic arguments whose charges I have aimed to dispel.

5. POSTSCRIPT: PROCESSES FOR EMPIRICISTS

I have described causal processes as systems of continuously manifesting relations between objects with causal properties and concomitant dispositions. This is to use ‘causal process’ as a term of art, for there are several established uses of the term. Especially relevant here, Salmon too thinks that the traditional analysis of causation in terms of relations between events is ‘profoundly mistaken’ (1984, p. 138), and he and Dowe both offer process theory alternatives. Their concepts of process, however, are different and more specific (arguably most relevant to causation in physics) than the one offered here. In this brief postscript, I aim to clarify the different questions to which these process theories are addressed. Salmon and Dowe give slightly diverging accounts, but as the differences are immaterial to the present discussion I will speak here of Salmon-Dowe (SD) processes. The SD account is tailor-made for an empiricist approach to causation, thus it should be no surprise that it differs from the one described here. Nevertheless, I will suggest that the proposal of Section 4 happily incorporates SD processes as part of the larger realist picture of causation.

The SD replacement for an analysis in terms of events is based on two fundamental causal concepts: process, and interaction. Salmon’s concept of process is inspired by Russell’s (1948, p. 477) notion of a ‘causal line’, denoting the persistence of an object: ‘Throughout a given causal line, there may be constancy of quality, constancy of structure, or a gradual change of either, but not sudden changes of any considerable magnitude’. This is vague, but Salmon goes on to describe processes more precisely. In his most considered view, shaped in large measure by Dowe, a process is something that carries or transmits a conserved quantity such as mass-energy, linear momentum, angular momentum, or electric charge. Causal processes propagate causal influences. They are altered in causal interactions, where conserved quantities are exchanged between processes.
Interactions occur only when causal processes intersect one another spatio-temporally, and though such intersections are necessary, they are not sufficient conditions for causal interaction. Interactions are the means by which modifications to causal processes are realized, and processes are the means by which the conserved quantities whose exchange constitutes an interaction are transported from one space-time location to another.

The SD account can thus be summarized in terms of the following definitions:

SD1: A causal process is a world line of an object that transmits (carries) or possesses (instantiates) a non-zero amount of a conserved quantity at each space-time point of its history.\(^{13}\)

SD2: A causal interaction is an intersection of world lines at which a conserved quantity is exchanged.

While the SD account views causation as an objective or mind-independent phenomenon, it is silent on the issue of \textit{de re} necessity. In this respect it is Humean. As I have noted, the intersection of two or more SD processes is a necessary but not a sufficient condition for causal interaction. Some processes pass through one another without interacting at all. Why then, one might ask, do some intersections yield interactions, while others do not? In true Humean spirit, we cannot say; there are merely regularities. Conversely, the realist hopes to explain why interactions take place on some occasions and not others. On the proposal of Section 4, the realist has a framework for such explanations. For on this view, something can be said about the sufficiency conditions for causal interactions: causation takes place whenever dispositions conferred by the causal properties of objects are manifesting; causal interactions simply are dispositions being manifested. Now as a good empiricist, Salmon disavows ‘powers’. Causal processes, he says (1984, pp. 202–203), simply ‘carry’ probability distributions for different interactions, and he calls these probabilities ‘propensities’. There is a perfectly Humean reading of probability in terms of frequency, of course, but the term ‘propensity’ does rather suggest a disposition, which is precisely what Salmon denies.

Both Salmon and Dowe maintain that their accounts shed light on the connection between causes and effects. As Salmon (1984, p. 155) puts it: ‘The propagation of causal influence by means of causal processes constitutes, I believe, the mysterious connection between cause and effect which Hume sought’. But SD processes do not constitute Humean causal connections. Consider the
following sequence of events: I strike the white ball with a cue, the white ball moves in the direction of the black ball, the white collides with the black, the black moves on. Let us idealize the example so that the motion of the white ball strictly satisfies SD1; that is, let us assume that it exchanges no linear momentum with the table or surrounding air molecules en route to the black ball, thus constituting an SD process between the interactions with the cue and the black. It is immediately apparent that this is not what Hume has in mind when he speaks of causal connections. For Hume, the striking of the white ball, the subsequent motion of white, the collision with the black ball, and subsequent motion of the black are all events. Hume sought necessary connections between these events, and found none. By simply re-labelling some of these events ‘processes’ as described in SD1 and calling them causal connections, Salmon misrepresents the very nature of the connection Hume dismissed.

Dowe also holds that his process view answers the question of what connects causes and effects, though he does not claim to be answering Hume per se. He describes causal connections as follows (2000, p. 171): ‘Interactions $I_1, I_2$ are linked by a causal connection by virtue of a causal process $p$ only if some conserved quantity exchanged in $I_2$ is also exchanged in $I_1$, and possessed by $p$’. Causes and effects are understood here as events or facts, which ‘involve the possession, or change in value of, some conserved quantity’. There may be an equivocation here, for although in his descriptions of causal connections Dowe identifies causes and effects with interactions (cf. SD2), his clarification of them appears to be consistent with causes and effects being processes (cf. SD1). In any case, the moral here is the same as that derived above in the context of Salmon. The SD view does not furnish an account of what Hume would call necessary connection, nor does it offer a framework for explanation with which to facilitate an account of the sort of de re necessity causal realists are after. The SD view is a process theory for empiricists and as such, the motivations of causal realism are foreign to it.

This is not to say, however, that causal realists should be hostile to SD processes – on the contrary! In a large range of cases, particularly in physics, the kinds of properties most relevant to the processes described in Section 4 are those concerning conserved quantities. Having determinate values of energy, momentum, charge, etc. confers dispositions on the objects that have them, and it is the continuous manifestation of these dispositions in consort with the presence and absence of other objects and their properties that
constitutes realist causal processes. To put it crudely, descriptions of SD processes cover part of what is happening in these cases, and the realist augments these descriptions by explaining how and why.

This subsumption of the SD account typifies the realist’s attitude towards any theory that aims to explicate causation in non-causal terms. Ehring (1997), for example, describes causation in terms of the activities of tropes: their persistence, destruction, fission, fusion, transference, and exchange. But trope theory is an account of properties, and so trope causation is further grist to the realist mill. The situation is not all irenic, however, because this subsumption is asymmetrical. Realists may absorb empiricist analyses of causation, absent the view that there is nothing more to say. But empiricists will always oppose the primitive causal concept of a power or disposition, so central to the proposal for causal realism given here, in response to the incoherence arguments with which we began.

ACKNOWLEDGEMENTS

I would like to thank Thomas Dixon, Katherine Hawley, Peter Lipton, Jack MacIntosh, and Hugh Mellor for extremely helpful discussions, two referees for this journal for excellent suggestions, and SSHRC for financial support. Earlier drafts benefited greatly from interactions with audiences at the annual conferences of the British Society for the Philosophy of Science and the Canadian Society for History and Philosophy of Science, and with colloquium participants at the Universities of Calgary and Leeds.

NOTES

1 For an overview of causal realism in similar terms, see Costa (1989, pp. 172–4).
2 This is the standard exegesis, though some argue for the surprising thesis that Hume was a causal realist; see Craig (1987, 2000) and Strawson (1989).
3 Fales ultimately agrees that we do not perceive causal necessity (p. 317, fn. 26), but for the wrong reason, I think: viz., that we never have impressions of the totality of causally sufficient conditions for effects.
4 Anscombe (1993/1971) comes closest to denying this: she holds that the causal relation is observable, but does not think that it is necessary.
6 For example, see Horwich (1987, pp. 170–172) for criticisms of Lewis’ ban on ‘back-tracking’ counterfactuals.
Blackburn (1993, p. 103) argues this as an exegesis of Hume.

This sort of approach is taken, e.g., by Mellor (1995, pp. 54–56, 219–220). Strictly speaking Mellor argues that ‘facta’, not events, are the relata of causation, but the approach is the same.

See also Fales (1990, p. 131).


Salmon first described causal processes as ones that are capable of transmitting a mark: a signal, information, or modification in structure (summarized in (1998, pp. 193–199)). He abandoned the mark criterion in response to Dowe (1992), who introduced the conserved quantity view. Salmon’s later position that processes transmit invariant quantities appears in (1998/1994), and his acceptance of conserved quantities comes in (1997).

To be precise, Salmon holds that the object must transmit, and Dowe that it need only possess, a conserved quantity. Only Salmon gives the explicit qualification of a ‘non-zero amount’. It is arguable that these differences are not substantive, but I will not consider the matter here.

Cf. Fales (1997, p. 122): ‘The proper “level” for causation, and hence causal processes, is the level of properties’.

REFERENCES


IHPST, Victoria College
University of Toronto,
91, Charles Street West
Toronto, Ontario
Canada M5S 1K7
E-mail: anjan.chakravarty@utoronto.ca

Manuscript submitted 30 June 2004
Final version received 23 March 2005