CAUSALITY & INTRINSICALITY

I

What is the relationship between causality and intrinsicality? Is causation intrinsic? And what would it mean to claim this? Such questions have recently received some, much deserved, attention. Not only is there something strikingly intuitive in the idea that what causes what is determined by factors intrinsic to the causal process; its recent employment in dealing with tricky late preemption cases offers a line of argument for some connection.¹

Let’s begin by taking one such case: Killer and Backup are lying await, ready to shoot Target. They both take aim, but Killer’s bullet gets there first. Let’s stipulate that Backup’s bullet is only thwarted by Target’s death. So there is no point, in the events leading up to the effect, at which there fails to be a back-up that suffices for the effect.

Cases of late preemption cause problems for numerous theories of causation. For instance, the popular counterfactual analysis is found wanting.² This states that C is causally dependent upon E iff C and E are distinct events and if C had not occurred, then E would not have occurred. C is then said to cause E iff there exists a chain of causally dependent events linking C with E.³ This, however, won’t do. There is no event in the causal chain initiated by Killer that is such that, if it hadn’t happened, the effect wouldn’t have occurred, for Backup’s bullet is always just right behind.

A natural reaction to these kinds of cases is to suggest that the analysis goes awry because it doesn’t capture the intrinsicality of causation. It allows that things extraneous to the causal relation can make a difference to whether C causes E. If we remove the background noise created by the back-ups in preemption cases, and focus just on the causal relation between C and E, then our theory of causation should give us the right results.

¹ See, for instance, Lewis (1986), Hall (2004a & b) and Menzies (1996).
² Whilst the problem of late preemption extends to other (non-process) reductive analyses, such as nomological and probabilistic approaches, to make the discussion manageable I shall focus on the counterfactual analysis.
³ See Lewis (1986).
This reasoning underlies one of Lewis’s amendments to his original counterfactual analysis. According to this amendment, roughly, C causally depends upon E iff E counterfactually depends upon C, or the majority of E’s duplicates counterfactually depend upon C.\textsuperscript{4} The guiding idea is that, since late preemption cases are unusual, in most instances where there are duplicates of C and E, there won’t be any rival backups to spoil the counterfactual dependence between them. How does the solution unfurl in the case of Killer and Target? If we duplicate the series of events from Killer’s shot to Target’s death, but remove extraneous factors such as Backup’s shot, then Target’s death is counterfactually dependent upon Killer’s actions. So he is rightly labelled the murderer. But the same isn’t true of Backup. In counterfactual situations where Killer’s shot is removed, there is a complete chain of counterfactually dependent events running from Backup’s shot to Target death. But this process or chain of events is different to that which occurs in the actual world, since in the actual world, Killer’s shot creates a spatiotemporal gap between the trajectory of Backup’s bullet and Target’s wounds. Since there is this failure of match between the actual world and one where Killer’s shot is absent, Backup’s actions don’t cause Target’s death.

Lewis has since rejected this amendment, for reasons that we shall come to. Hall, however, argues that we are right to think that there is a close relationship between causality and intrinsicality. To be more precise, he argues that this Intrinsicality Thesis holds true of one kind of causation – production:\textsuperscript{5}

Let S be a structure of events consisting of event e, together with all of its causes back to some earlier time t. Let S’ be a structure of events that intrinsically matches S in relevant respects, and that exists in a world with the same laws. Let e’ be the event in S’ that corresponds to e in S. Let c be some event in S distinct from e, and let c’ be the event in S’ that corresponds to c. Then c’ is a cause of e’ (2004b: 264).

A key notion here is that of ‘a structure of events’. These structures, Hall writes, are built up out of events, and their intrinsic natures are exhausted ‘by the intrinsic natures of their constituent events, together with facts about how those events are spatiotemporally juxtaposed with one another’ (2004b: 262). According to Hall’s

\textsuperscript{4} See Lewis 1986: 206.
\textsuperscript{5} Hall (2004a) thinks that we are pushed to accept that there are two varieties of causation – production and dependence – because a number of plausible theses about causation, of which the Intrinsicality Thesis is one, cannot be true of every kind of event causation.
Intrinsicality Thesis (HIT), altering the environment of an event structure, whilst holding that structure fixed, cannot alter the fact that the events in the structure are causes of e. So, roughly speaking, if c is a cause of e, and the laws are held fixed, then any duplicate of c is also a cause of e.

Is HIT true, and does it capture what we want to say about the relationship (if there is any) between causality and intrinsicality? Sections II and III argue that some proposed counterexamples to HIT, or a close cousin of HIT, are inconclusive. Section IV discusses another counterexample which, whilst again not decisive, raises more serious doubts. I end by considering an alternative intrinsicality thesis.

II

One reason for rejecting HIT, offered by Lewis and others, is that instances of double prevention seem to constitute counterexamples. Let’s return to the scene of Target’s murder, but now imagine that Boss has become wise to Killer’s plans. She decides to send Guard to shoot Killer. However, Backup, wanting Target dead, intercepts Guard and murders him before he has a chance to kill Killer. Here we have a case of double prevention: Backup’s actions prevent Killer’s death which, in turn, prevents Target’s life from being saved.

Intuition, the objection goes, tells us that Backup caused Target’s death, since if it wasn’t for Backup’s actions, Target would have survived. But there is no intrinsic process running from Backup’s actions to Target’s death, for we can alter the causal status of Backup’s action simply by altering something extrinsic to the process. For instance, we could suppose that Guard does not receive orders from Boss to shoot Killer, so Backup does not prevent Killer’s death. To put it in Hall’s terminology, there are possible worlds which duplicate the event structure consisting of Backup’s killing of Guard, to Guard’s failure to shoot Killer and Killer’s killing of Target, but Backup isn’t the cause of Target’s death in those worlds. Since, in those worlds, Guard poses no threat to Killer.

Does HIT stand refuted? I think we need to be cautious in forming a judgement at this point, since intuitions regarding these kinds of cases vary depending on how

---

6 See Lewis (2004a), McDermott (1995) and Schaffer (2000b). Hall (2004a) argues that this is a reason to think that we have two concepts of causation.
they are described. Suppose, for instance, that Backup isn’t on the scene but that Guard’s kindly Aunt Agnes delays him with another cup of tea, giving Killer time to complete his dirty deed. Do we want to claim that Aunt Agnes caused Target’s death? The intuition that Backup did cause Target’s death, whilst far from being universally shared, is felt much less keenly here. But the structure of the situation is exactly the same, so what should we conclude? I think that we should say that these cases are too controversial to be the building blocks for a theory of causation. Given that the cases involve omissions, which are notoriously tricky, our response to cases of double prevention should arise from independently motivated theories of causation, rather than be the data upon which those theories are built.

This gives proponents of HIT more room to manoeuvre. Broadly speaking, two strategies can be employed to defend HIT. These are disparate, but either can be motivated given certain analyses of causation. The first response is to simply deny that Backup, or Aunt Agnes, is the cause of Target’s death. Probably the most promising justification of this claim stems from the idea that omissions aren’t causes, so Guard’s failure to kill Killer is not a cause. This initially rather counterintuitive sounding thesis tends to fall out of process theories of causation. So if these analyses can be independently well-motivated, there’s no reason to reject HIT on the basis of double prevention cases. Another possibility is to deny transitivity, arguing that whilst Backup preserves Killer’s health which causes Target’s death, nevertheless, Backup isn’t the cause of the death. However, some reason for believing this still needs to be found. Perhaps a more plausible justification for denying that Backup or Aunt Agnes causes Target’s death invokes the distinction between causes and conditions. We could argue that Backup’s actions are required to preserve one of the conditions of the cause, namely Killer’s health. However, they do not cause Target’s death, as more stringent requirements than counterfactual dependence need to be in place for her actions to count as a cause.

But what if proponents of HIT want to maintain that omissions can be causes, that causation is transitive and that counterfactual dependence is sufficient for causation? Then they might adopt the second strategy. This allows that Backup or

---

7 For discussion see, for instance, Dowe (2000) and Lewis (2004b).
8 Or, alternatively, we could say, with Hall (2004a, §6), that they don’t cause them in the full-bloodied, ‘production’ sense.
9 See, for instance, Dowe (2000).
Aunt Agnes causes Target’s death, but then claims that these cases are misdescribed. Properly understood, they aren’t counterexamples to HIT. The key question here is: which events should be incorporated into the causal structure of Target’s death? Defenders of HIT can argue that if we simply think of the causal structure of Target’s death as running from Backup’s shooting of Guard to Killer’s health, then finally to Target’s murder, we have missed out an essential part of the event structure. For Backup or Aunt Agnes only preserve Killer’s health given that Guard poses a threat. So the event of Guard receiving Boss’s orders is part of the event structure for this cause. Now we no longer have a counterexample to HIT, as the causal status of Backup’s (or Aunt Agnes’) action does not vary in the worlds where we duplicate the suggested alternative event structure.

Hall objects to this response on the grounds that it renders too many events causes of Target’s death. It is just counterintuitive, he claims, to think that one of the causes of Target’s death are Boss’s orders to Guard. I agree that this does sound counterintuitive, but I am not convinced that this need worry defenders of HIT, granted the theoretical assumptions already in play. For once we have accepted that omissions can be causes, transitivity and the sufficiency of counterfactual dependence for causation (the theses required to block the first strategy), very many counterintuitive things are going to have to count as causes. Not only will Aunt Agnes be a cause of Target’s death, but also the Queen’s failing to water my flowers will be a cause of their death, my birth will cause my death and, to borrow Bennett’s example, the April showers (by preserving the forest from the May lightening) will cause the forest fires in August. Whilst it is true that we don’t tend to pick these things out as the cause or even as a cause, if we adopt causation via omissions, transitivity and the sufficiency of counterfactual dependence, this will be due to contextual and pragmatic factors concerning what is salient to the inquiry.

---

10 See Hall 2004a: 245.

11 See Bennett 1987: 373. See also Lewis 2004a. Lewis’ discussion of the Red and Black cases suggests that he should be willing to accept that Boss’s orders are a cause of Target’s death.

12 We might argue that only some omissions that stand in the relation of counterfactual dependence are causes. For instance, perhaps only those omissions that form part of the causal structure of an event (where our choice of omissions within that structure reflects the possibilities we are willing to take seriously) count as causes (see Woodward 2004 §2.8). Hence we could maintain that the Queen’s failure to water my flowers is not a cause, whilst allowing that the Guard’s failure to shoot Killer is. This position cannot be discussed here. But it does commit us to the view that ‘conclusions about causal relationships are sensitive to one’s choice of representation’ or causal structure (Woodward 2004: 80). Since the question, ‘which events form part of the causal structure?’, can only be answered by
So if we deny that omissions are causes, that transitivity holds or that counterfactual dependence is sufficient for causation, then a defender of HIT can simply reject the supposed counterexample. However, if these background causal assumptions are embraced, then the causal structure of any event will be very rich indeed. Defenders of HIT can say that Boss’s orders to Guard are part of the event structure since, if it is excised in duplicate situations, Backup is no longer the cause of Target death. Of course, opponents of HIT may object that this begs the question against them, since they are arguing that events extrinsic to the event structure, such as the fact that Guard received the orders he did, can affect the causal status of Backup’s actions, so HIT is false. What we need to break the deadlock is some principled answer to what should and shouldn’t count as part of the event structure. The only criterion that is offered by Hall is that an event e counts as part of the event structure of an event e* if it is numbered amongst its causes, but this clearly fails to decide the issue either way.\(^{13}\)

Where does this leave us? In the absence of an answer to which events count as part of the event structure, we can justifiably conclude that the argument from double prevention is inconclusive. For whether or not we should accept the description of the cases given by opponents of HIT, will crucially turn upon whether or not we already have accepted the intrinsicality thesis.\(^{14}\)

III

Another class of potential counterexamples to HIT invokes instances of indeterministic causation.\(^{15}\) To illustrate, suppose that an unstable radioactive atom, b, has a small chance of decay, say 0.1, determined by its half life. But if it is bombarded by a photon, then the chance that b decays rises to 0.9. Now suppose that, in the actual world, an event of photon bombardment (e1) takes place and this causes the decay of b employing causal information that cannot be analysed reductively, it offers a very different picture from the Lewisian framework assumed by many participants of this debate.\(^{13}\)

Due to time constraints, I shall not talk about the problems that omissions cause for HIT. However, I think that if we can circumvent the problem of locating omissions, then the same pair of responses can be given. In other words, either we could deny that omissions are causes properly so-called, or we could say that some omissions do affect the intrinsic causal structure of an event and so count as part of the duplicated situation. See the comments on locating omissions in sections IV and V.\(^{13}\)

Variations of this kind of case have been offered as arguments against a reductive analysis of causation (see, for instance, Tooley 2004).
(e2). If we duplicate the structure of events, so we replicate the event of b’s bombardment, the decay of b and ‘facts about how those events are spatiotemporally juxtaposed with one another,’¹⁶ then there will be nomologically possible duplicate situations where the photon’s bombardment isn’t the cause of the atom’s decay. For it is nomologically possible that e1 might stand in the same spatiotemporal relations to event e2, but e1 not be the cause of e2, since e2 occurs spontaneously. So the relation between e1 and e2 in the actual world will fail to satisfy HIT.

How should proponents of HIT respond? The first strategy is to deny that e1 causes (or produces) e2. There is plenty of room for disagreement here. Perhaps we should say that there is no definite fact of the matter concerning whether or not e1 caused e2 in the actual world (and in nomologically possible worlds where this sequence of events is duplicated). It is not simply that we can’t tell whether or not e1 is the cause. It is nomologically possible that e1 caused e2 and nomologically possible that it didn’t, and there is simply no fact about which is the case in the actual world. If this is so, then HIT is fully vindicated, since this indeterminacy will be duplicated across nomologically possible worlds. This response, however, may well strike us as ad hoc: if there is a determinate fact of the matter concerning whether or not e2 is self-caused in the absence of e1, why must this change just because e1 comes into spatiotemporal view?

If a defender of HIT wants to maintain that the photon’s bombardment is the cause of the atom’s decay, this leaves the second strategy: that of denying that this is a counterexample to HIT, by claiming that there are no relevant worlds where the duplicate of e1 fails to be the cause of the duplicate of e2. One response that can put aside is this: if we allow that causation is an irreducible relation, then the counterexample doesn’t work, since this special irreducible relation becomes part of what is duplicated between e1 and e2. This, however, is something that Hall himself wants to reject.¹⁷ More importantly, it begs the question to assume that if we duplicate the structure of events between e1 and e2, this will include the irreducible relation of causation. For, in this context, we are considering whether or not HIT is true. So we cannot claim that this irreducible relation is something which will be present in all the relevant duplicate situations, without already having assumed the truth of HIT.

¹⁶ Hall 2004b: 262.
A more promising justification for the claim that there are no worlds where a duplicate of e₁ fails to cause a duplicate of e₂, begins with the popular idea that causes raise the chance of their effects. Lewis writes,

If distinct events c and e both occur, and if the actual chance of e (at a time immediately after c) is sufficiently greater than the counterfactual chance of e without c, this implies outright that c is a cause of e (1986:180).

On this view, it looks like we can simply deny the supposed counterexample to HIT. For all the nomologically possible worlds which contain duplicates of e₁, e₂ and the spatiotemporal relations between them, are worlds where e₁ raises the probability of e₂ (at least if we ignore those pesky worlds with backups around to spoil the probabilistic dependence). So e₁ is the cause of e₂ in those worlds.

Many will feel resistant to this: surely it is nomologically possible to have e₁, e₂ and certain spatiotemporal relations pertaining between them, but e₂ not be caused by e₁? What seems to be a nomological possibility has been ruled out a priori by our conceptual analysis of causation. Lewis, however, defends this counterintuitive result by utilising the compelling thesis that truth supervenes upon being.¹⁸ We are tempted to think that there is a nomologically possible world where, despite the presence of both e₁ and e₂, e₁ isn’t the cause of e₂, for we wrongly suppose that there is a nomologically possible world where the counterfactual, ‘if e₁ hadn’t occurred, e₂ would have anyway’ is true. But there are no nomologically possible worlds where there is a fact of the matter upon which the truth of this counterfactual can supervene. Since we can never establish counterfactual dependence in indeterministic settings, we have to settle for causation as probabilistic, rather than counterfactual, dependence.

This is a little quick, however. There may be something for the truth of the counterfactual to supervene upon, something that is disregarded by HIT and the probability raising account. For suppose that there are some physical relations, not captured by the description of the intrinsic properties of the events in the structure and their spatiotemporal relations. For instance, perhaps causation requires BIFF, a placeholder for something like property or energy transference, and this involves the persistence of an entity, a persistence which cannot be analysed in terms of spatiotemporal or nomological relations between temporal parts.¹⁹ Given this, there

---

¹⁹ For a more detailed description of such a view, see Ehring (1997).
could be something for the truth ‘e1 doesn’t cause e2’ to supervene upon. For it may be the case that whenever e1 causes e2, there is this transference of energy which is absent in some cases where duplicates of e1 and e2 stand in the right spatiotemporal and probabilistic relations. So HIT could be false, but the causal truth could still have some being upon which to supervene.

Suppose, however, that we reject Hall’s claim that the only relevant relations that form part of the event structure are spatiotemporal ones. This seems a little arbitrary in any case: if there are other, non-spatiotemporal relations, that can exist unaccompanied by all other entities distinct from that event structure, why shouldn’t they form part of the duplicate situations? Once these are included, this more permissive form of HIT, call it HIT*, is only vulnerable to the indeterministic case offered on the assumption that it is nomologically possible to duplicate e1, e2 and all of the relevant relations, and yet e1 fail to cause e2. In this case, if we do not postulate an irreducible relation of causation, Lewis’s point does hold, since we have a true proposition, namely that e1 doesn’t cause e2, which fails to supervene on any being.

The proposed counterexample to HIT* accordingly becomes even more controversial. Is it really the case that we could duplicate absolutely everything about the scenario and yet e1 fail to cause e2? At this point, the lesson of the last section can be applied. Given the controversial nature of the thought experiment, if there is good reason to suppose that causation can be identified with B IFF, or probability-raising, then we seem well within our rights to simply reject the proposed counterexample.

IV

Instances of trumping offer yet another challenge to HIT. To illustrate, suppose that all particles that are subject to a Y field accelerate along a curved trajectory. According to HIT, since the Y field is a cause of the particle’s curved trajectory in the actual world, all the duplicates of this event structure should also cause this movement. But now suppose that there is a law stating that if a particle is subject to both X and Y fields, the X field alone causes the resultant curved trajectory. In this case, in duplicate situations where the X field is present, the duplicate particle will move because of the

\[20\]

The following example is a simplification of a trumping case offered by Schaffer, 2000a: 173-4.
presence of the X field. We thus have another counterexample to HIT. For in a
counterfactual world which precisely duplicates the Y field and movement of the
particle, but which introduces the X field, the Y field doesn’t cause the particle’s
curved trajectory. So the extrinsic presence of the X field affects whether or not the Y
field causes the particle’s curved trajectory.

How can proponents of HIT respond? The same pair of replies utilised in cases
of double prevention and indeterministic causation can also be employed here. But
they appear less successful in this context. Let’s start with the first strategy. We could
simply deny that the Y field causes the curved trajectory in the actual world, or
perhaps deny that it is an instance of full-blooded production. But this reply seems far
less attractive in this case as compared to those involving omissions. Whilst we can
easily offer examples that make causation by omissions look rather suspect, there
seems no reason to think that this case isn’t a paradigmatic example of causation. The
Y field does produce the particle’s curved trajectory in the actual world.

What, then, of the second strategy? We could deny that this is a
counterexample to HIT, by denying that the counterfactual situation where both the X
and Y field are present is a duplicate of the actual situation in all of its relevant
respects. For, we could say, the presence of the X field alters the character of the
situation. Once again, then, we meet the problem of determining the event structure for
an effect, and hence the relevant duplicate situations. In the double prevention case
offered, however, we could at least trace a causal chain from Boss’s orders to
Backup’s killing of Guard and Target’s eventually murder.\(^{21}\) And, given the
assumptions about causation mentioned earlier, there is reason, independent of HIT, to
accept that an event which attempts to prevent a certain effect can also ultimately be
amongst it causes.\(^{22}\) In our trumping case, in contrast, no such theoretical justification
seems to be in the offing. The Y field is not part of the causal chain between the X
field and the particle’s curved trajectory. If we suppose further that the Y field has no
influence at all upon that process, it does seem something entirely extraneous to it.

Even if we put this worry aside, however, the response is still problematic,
since it does nothing to restore the proposed motivation for HIT as an answer to
preemption problems. To illustrate, let’s again assume the counterfactual analysis of

\(^{21}\) At least, we could if we allow that omissions can be causes – an assumption that will be
endorsed by those pursuing this response.

\(^{22}\) See, for instance, Lewis (2004a).
causation, and suppose that we are now in the counterfactual situation where both the X and Y fields are present. Here, we want to say that the X field is the cause of the particle’s trajectory. So, according to the counterfactual analysis, we must look to see whether the X field stands in the right counterfactual relations in duplicate situations where the background noise is removed. But if the presence of the X field changes the intrinsic structure of the process in the counterfactual, as compared to the actual situation, for this is why the Y field causes the curved trajectory in the actual world, then the presence of the Y field will also count as a relevant difference between these two counterfactual scenarios. As a result, all the relevant duplicate situations of this scenario will contain a Y field as well as an X field, thus destroying the counterfactual dependence between the X field and the particle’s curved trajectory. Unless we can find some asymmetry here then, the X field won’t count as the cause of the particle’s trajectory in this counterfactual situation.

Perhaps, however, we just need to look harder for the source of this asymmetry. We could argue that the reason why the counterfactual situation, where both the X and Y field is present, is not a duplicate of the actual situation in all of its relevant respects is because, in the actual world where we have just the Y field causing the particle’s curved trajectory, the absence of the X field is part of that process. Defending this view commits us to the claim that in the counterfactual situation, there is something missing, the thing that makes the negative existential statement, ‘there is no X field’, true. It won’t do, for instance, to say that the failure of the X field to occur just supervenes upon the positive facts of this situation. For these positive facts – the Y field, the intrinsic (non-causal) relations, and the event of the particle’s curved trajectory - can be precisely duplicated as a part of the situation where there is the Y field and the X field causing the particle’s trajectory. So the X field would count as something extrinsic, altering the causal status of the Y field. It must be the case that the absence of the X field changes the intrinsic character of the event structure in the actual situation; in other words, there must be something extra, namely an absence, which is missing in the trumped case. For only then will we have the relevant asymmetry between the actual and counterfactual situations.

23 Cases of trumping, however, cause problems for process, as well as nomological and probability, approaches to causation.
24 Tooley (2004) takes this line.
This seems to be the only way of defending HIT against these trumping cases, granted their coherence. However, it comes with a heavy burden – that of treating omissions, such as the failure of the Y field to occur, as locatable entities that form part of the process or event structure. Hall argues that omissions cannot be located. Whether or not he is right about this is a difficult issue that cannot be decided here. If we are sceptical of this claim, however, then defenders of HIT have to contest these trumping cases directly.

For trumping cases to work, we must suppose that the trumped line runs to completion but, for some reason, this alternative line just ‘doesn’t count’. We can’t say that the trumped line doesn’t count because there is some mechanism in place that blocks its efficacy. For, then, trumping cases can simply be treated as instances of preemption, so our counterexample and pairing problems disappear. But, then, what does this ‘doesn’t count’ actually amount to? Without an answer to this, shouldn’t the case be described as an instance of overdetermination rather than preemption? For if the process between the Y field and the particle’s curved trajectory really is the same, what reason is there to recognise the existence of a law which states that if a particle is subject to both X and Y fields, the X field alone causes the resultant trajectory, rather than both the X and Y field?

These cases certainly do seem underspecified at the crucial juncture. Perhaps there is always some mechanism that blocks the effectiveness of the trumped line. Or, if there isn’t, we should be willing to describe this as a case of overdetermination, rather than preemption. But, we might argue, given that we are dealing with causation, not just as an empirical concept, but also as a folk psychological concept, surely the onus of proof is on those who deny the coherence of trumping cases? For there seems nothing incoherent in supposing that there is a possible world where the event structure between the Y field and the accelerating particle is replicated (bar its causal status) when the X field is present. Nor in the idea that, in such a case, there could be a genuine law which gives causal priority to the X field when both X and Y fields are present. And, if this is right, we do have a counterexample to HIT.

I agree that the burden of proof does lie with a defender of HIT here. But there is reason to think that it could be shouldered. If causation can be identified with BIFF,

---

25 See Hall 2004a and b. Of course, even if Hall is wrong on this point, this alone wouldn’t suffice to show that the absence of the X field is part of the cause in this case.

26 See Lewis (2004a) and Schaffer (2000a).
for instance, then these trumping cases would have to be rejected. Once more, then, a
defence of HIT could piggy-back on the success of a well motivated theory of
causation that endorses it. But does a defence of the intrinsicality of causation have to
await a defence of an analysis of causation that excludes trumping? Insofar as there is
something intuitive in the thought that, even if we grant the coherence of trumping
cases, the process between the Y field and the particle’s curved trajectory should count
as intrinsic, we might suspect not. Perhaps the problem is that HIT employs the wrong
conception of intrinsicality.

V

It has been noted, in relation to the more straightforward monadic properties, that the
duplication analysis, employed by HIT, doesn’t capture everything about intrinsicality.
Dunn, for instance, argues that α’s property of being identical to α is, intuitively, an
intrinsic property, as it is one that ‘the object has in virtue of itself, depending on no
other thing’ (1990: 178). But it doesn’t count as an intrinsic property on the
duplication analysis, since it isn’t a property which is had by α’s duplicates. One
response to this is to argue that the duplication analysis is false. But it is better, I think,
to follow Moore and allow that there are different senses of ‘intrinsic’. One sense of
‘intrinsically different’, Moore claims, refers to all those entities that are not
qualitatively alike.27 This is tracked by the duplication analysis, for by claiming that an
object’s intrinsic properties are those that are instantiated by its duplicate, it must be
possible for intrinsic properties to be had by more than one particular. Moore’s second
sense of ‘intrinsically different’, in contrast, concentrates on differences in the
‘constituents’ of a particular. Any differences in the entities that lie entirely within the
confines of the particular, count as differences in its intrinsic nature. Humberstone’s
notion of an intrinsic or ‘interior’ property expresses this idea. These are properties
whose existence and nature has been ‘entirely determined by what is the case within
the confines of the would-be possessor’ (1996: 242). So no entity distinct from its
instantiator can affect the property in any way.

How could we utilise this to modify HIT? The thought is that what matters for
intrinsicality when it comes to causation, isn’t that all the duplicates of this event

27 See Moore 1922: 262.
structure share its causal status, but rather that the causal character of the event structure, i.e. the facts about what causes what in that structure, depend only on the intrinsic properties of the events and those relations between the events that are ‘intrinsic to its relata’ and ‘intrinsic to its pairs’. For if the causal character of the structure is determined by those entities that lie within the confines of that structure, this captures the intuitive thought that when one thing causes another, there is no need to go beyond that spatiotemporal region in order to discover what is responsible for this fact.

It might be thought that this alternative conception of intrinsicality has the resources for dealing with trumping cases. If, when the causal relation between the Y field and the particle’s curved trajectory does obtain, this fact is determined solely by those intrinsic properties of the relata and those relations that only require the presence of those relata, then surely this should count as an interior relation? For, granted we have rejected the idea of omissions as locatable entities, the failure of the X field to occur isn’t any thing that the causal relation can be determined by. But what does it mean to say that these entities solely ‘determine’ the causal character of the relation or event structure? And does this suffice to give us an adequate notion of intrinsicality?

Vallentyne offers an analysis of intrinsicality which purports to capture this wider notion of intrinsicality as interiority. He begins by utilising the notion of ‘a contraction of a world’ (1997: 209). This is one which is obtainable from the original world by subtracting objects from it. Whilst a ‘maximal contraction’ of a world (‘an x-t contraction’) is a world obtainable from the original one by removing all objects wholly distinct from x, and all spatiotemporal points not occupied by x. An intrinsic property, Px is then defined as (a) one that remains Px in each x-t, and (b) likewise for not-P. This is not a reductive definition. But if we agree with Vallentyne that the notion of a contraction is ‘intuitively clear and familiar’ (p.213), it may be helpful to extend his definition to the trickier arena of relations. A maximal xRy contraction can be thought of as one obtainable from the original world by removing all objects and spatiotemporal points unoccupied by xRy. An intrinsic relation can then be defined as

---

28 See Lewis 1983: 26. A relation is ‘intrinsic to its relata’ if it depends solely upon the nature of its relata, taken independently. A relation is ‘intrinsic to its pairs’ if it depends solely upon the relata and those relations that could hold between the relata, when they are unaccompanied by all other entities.

29 Vallentyne, 1997: 211.
one that remains (a) $xRy$ in each maximal $xRy$ contraction of the world, and (b) likewise for $\neg xRy$. So having the same teacher is an extrinsic relation, since $x$ and $y$ do not stand in that relation in contractions of the world which remove all objects and spatiotemporal points distinct from $xRy$. Whereas, being taller than is an intrinsic relation, for in each maximal $xRy$ contraction, $x$ will still stand in that relation to $y$.\(^{30}\)

With this definition in play, however, we seem to have enough to scupper the idea that causation could be an intrinsic, in the sense of interior, relation. The causal relation between the Y field and the particle’s curved trajectory has no difficulty meeting the first condition, since it will remain a cause in all maximal $xRy$ contractions of that world. However, now consider the world where we have both the X and Y field. According to the second condition, that the Y field doesn’t cause the particle’s curved trajectory should remain true in each maximal $xRy$ contraction of the world. But it doesn’t. In a contraction which removes the X field, the Y field will change from not being a cause to being a cause of the particle’s curved the trajectory. So the causal relation will count as extrinsic according to this definition.\(^{31}\)

We might argue against this conclusion by making the same move that was canvassed on behalf of HIT: we could claim that an essential part of the cause is missing from the purported maximal $xRy$ contraction of the world, namely, the absence of the X field. This is not in the contraction since we are considering a world in which the X and Y field occur together. As such, the Y field does not count as the cause of the particle’s curved trajectory in the maximal $xRy$ contraction under consideration. But, again, this requires us to locate omissions, so the move to the interior notion of intrinsicality has not bettered our position at all.

Should we just give up on the half of Vallentyne’s definition of intrinsicality which causes problems for us? Unfortunately not, since then properties such as ‘being

\(^{30}\) Unfortunately, it is much more difficult to envisage what these $xRy$ contractions involve in the case of relations, as compared to objects, since they lack such clearly defined boundaries. In order to capture relations that are not only intrinsic to their relata, but also intrinsic to their pairs, maximal $xRy$ contractions have to leave intact all those relations that only require the existence of the relata. Consequently, we may well wonder whether this talk of contractions offers any advancement on the original gloss. I think the definition has some worth, however, if only in that it gives some indication of what talk of ‘determination’ involves in this context.

\(^{31}\) Vallentyne states that these maximal contractions involve the removal of the laws of nature (thus rendering all causal relations extrinsic). This seems debatable (since the truthmakers for laws may be immanent universals) but, in any case, if this view is going to remain a close relative of HIT, we would need to stipulate that the laws are to be excluded from the maximal contractions.
the tallest person’ will end up being intrinsic, for an object that has that property would retain it in each x-t contraction. Vallentyne’s definition requires the second, negative, part so as to count such properties extrinsic; it is false that someone who lacks the property remains not-the-tallest-person in each x-t contraction of the world. At this point, I think it is safest to concede that if trumping cases are a genuine possibility, and omissions cannot be located (and treated as part of the causes in trumping cases), then causation doesn’t count as intrinsic, in the duplication or interior sense.\textsuperscript{32}

But if this is so, how can we account for the intuition that the causal relation between the Y field and the particle’s curved trajectory is intrinsic, even granted trumping? Consider the following principle: for any event structure, consisting of event e together with all its causes going back to some earlier time t, all the causes of event e will remain causes of event e given each maximal ES contraction of the world.\textsuperscript{33} Schaffer’s trumping cases do not challenge this first part of Vallentyne’s definition. What the trumping case shows is that the causal relation between the Y field and the particle’s curved trajectory can be prevented from holding due to something extraneous to that relation, namely the presence of the X field. But it does not show that the fact that the Y field caused the particle’s curved trajectory depends upon any thing more than the intrinsic properties of and intrinsic relations between these events, when combined with the negative existential fact that there is no X field. Moreover, whilst it is highly controversial what makes this negative existential statement true, it will remain true in each maximal ES contraction. So the removal of all objects and spatiotemporal points distinct from the event structure will not alter the fact that this Y field caused this particle’s curved trajectory.

It might be useful to use an analogy here. Sider (2001) argues that the property of being spherical is extrinsic, since the object could loose the property given certain accompaniments. If a spherical object is embedded in a cube, for instance, then it

\textsuperscript{32} Even if we are unhappy with this extended use of Vallentyne’s definition, and prefer the original gloss offered, there are still grounds to worry here. For it is not clear that it is really appropriate to say that the causal status of this relation is entirely determined by entities that lie within its boundaries, given that the existence of the causal relation between the Y field and the particle’s trajectory depends upon whether or not some other entity (the X field) exists. Of course, an answer to this question awaits an explication of what ‘determines’ amounts to in this context, but on the face of it at least, the claim looks problematic.

\textsuperscript{33} A maximal ES (event structure) contraction of the world is a contraction obtainable from the original world by removing all objects and spatiotemporal points distinct from the event structure in question, whilst leaving the laws of nature in tact.
would cease to be spherical. But, nevertheless, the property of being spherical still seems, in some sense, importantly different from straightforwardly extrinsic properties, such as being ten feet away from a bear. For, given the absence of the cube which the spherical object is embedded in, it seems intuitive that the thing’s having that property is determined by what is the case within that object. The same is true of the causation case. Given the absence of the X field, the Y field’s causing the particle’s curved trajectory is determined by what is case within the event structure.

It is this, I suggest, that accounts for the fact that, even if we find trumping a coherent possibility, there is intuitively something intrinsic going on in the Y field’s causing of the particle’s curved trajectory. The proposal doesn’t require us to locate absences, since it allows us to claim that negative existential statements, such as the absence of the Y field, can be made true just by the positive facts of the situation. So what negative existential statements there are only requires the addition of something extra, like the presence of the X field, to change. However, the suggestion that this is what explains the intuition of intrinsicality is a deflationary suggestion. Since what I have described is not really intrinsicality at all, in either the duplication or interiority sense.

To sum up, the most powerful argument considered against the intrinsicality of causation, in either the duplication or interiority sense, is the trumping case. Double prevention cases are inconclusive. Indeterminisitic cases can be rendered harmless given a minor amendment to HIT. But trumping cases require more significant moves to be made. Unless some reason is found for simply rejecting the examples, the defender of the intrinsicality of causation is forced to treat absences as part of the causes in trumping cases, which requires a solution to the problem of locating absences. Furthermore, the lingering suspicion that, even granting trumping cases, causal relations involve some notion of intrinsicality, can be explained by pointing out that causation satisfies the positive part of Vallentyne’s definition of interiority. To this extent, causation differs from other clearly extrinsic relations, such as having the same teacher, but it is still a far cry from its counting as an intrinsic relation.

34 Although it doesn’t commit us to this. It might be that something like Armstrong’s totality facts (2004 §2.4), for instance, are also required.

35 Many thanks to Joel Smith, Paul Noordhof, Helen Beebee and audiences in London, Bristol, Manchester and a Metaphysics in Science workshop in Nottingham.
References


