Causal Indiscernibility

Jeff Engelhardt

1

Many discussions in the philosophy of mind presuppose the following principle; call it the Principle of Causal Exhaustion[1]:

CE If ‘a’ and ‘b’ refer to occurrences that are causally indiscernible, then a = b.[2]

If CE is false, many arguments in these discussions are unsound and other points are confused. But before we decide CE’s truth value, we must clarify the relevant concept of causal indiscernibility. We shall see that in fact there are several distinct concepts in play here. They are usefully distinguished along dimensions of (i) concepts of causation, as in sections 2 and 3, and (ii) modal force, as in section 3. But first, we must make a few preliminary remarks about CE.

2 Preliminaries

Intuitively, a and b are causally indiscernible if they have the same causal profile by some force of necessity. Imagine that an occurrence’s causal profile is a pair (a triad, a tetrad) of lists, one for causes (causal alternatives[3]) and one for effects (effectual differences[4]). Include in the items populating each list whatever variables are needed to individuate causes, effects, etc.—spatiotemporal region (i.e. Bennett’s “zones”), substance, property, etc.

[1] See [Low00, Rit05], [Sos95, Kim97, Ben08, Pea79, Kim89, Stu05, Kim93b, Kim95, Kim98, LL87], [Pap93], [Dre89], [Mal68]

[2] I follow David Papineau here in using “occurrences” to remain neutral with respect to the causal relata. If it turns out that events are the causal relata, then the principle says that the causal indiscernibility of “two” events entails “their” indiscernibility (full stop); if facts are the causal relata, then CE says that the causal indiscernibility of “two” facts entails “their” indiscernibility (full stop). And so on.

[3] See [Hit93], [Hit95], [Hit96]

[4] See [VF80]

Two causal profiles share an item only if all the variables in the item are the same and only if it appears in the same (cause, effect…) list in each profile. If, for example, the same ⟨object, property, time⟩ triple appears in the "cause list" of two occurrences, then these two occurrences have the same Kim event as a cause in at least this one case, and so their profiles share an item. Using this imagery, we would say that two occurrences are causally indiscernible just in case identity is a bjiective function between "cause lists", "effect lists", etc. in the two profiles.

But this leaves the modal force of "indiscernibility" unspecified. Let us employ the scenic idiom of possible worlds to clarify it. We might fill out a profile with all of an occurrence's causes and effects in the nomically possible worlds, in all the physically possible worlds, in all the metaphysically possible worlds, and so on. If the lists of physically possible causes and effects are isomorphic, then the relevant occurrences are physically causally indiscernible. So too, mutatis mutandis, for nomic and metaphysical causal indiscernibility.

Rather than give preference to any of these as the appropriate modal force for CE, let us just note that "causal indiscernibility" embeds a modal operator; where Cxy = x and y have distinct causal profiles', we would render "x and y are causally indiscernible" as: $\neg\Diamond$Cxy. The force of the modal operator is multiply ambiguous, of course, and thus, "causal indiscernibility" is multiply ambiguous. It has as many readings as there are modal strengths. In turn, then, there should be just as many readings of CE. And thus, not all readings of CE range over all causal relations. As such, not all true claims of causal indiscernibility between a and b entail that a and b have all the same causes and effects. This point is obvious once it has been made, but it's worth being deliberate here; otherwise, it's tempting to think that only causally irrelevant features of an occurrence could falsify CE. But on the contrary, causal relations in metaphysically possible worlds may falsify the claim that, say, *nomic* causal indiscernibility suffices for identity. A great deal turns, then, on the strength of indiscernibility that the arguments and debates assuming CE can defend. Since CE arises most commonly in discussions of mental occurrences and causation, one looming question is, What is the largest stable set of possible worlds for which mental and physical occurrences share causal profiles?

Hence, not all strengths of CE range over all possible causal relations.

In addition, CE does not range over any occurrences that are outside the causal nexus in a given set of possible worlds. Let us say that an occurrence is "outside the causal nexus" just in case it stands in no causal relations whatsoever. Cartesian souls (in the nomically possible worlds), numbers (in the metaphysically possible worlds), times, facts, or properties may be outside the causal nexus. If there are any such occurrences, then they have empty causal profiles. That is, since such an occurrence stands in no causal relations, its list of causes is blank, as is its list of effects (alternative causes, effectual differences, etc.) Thus, all such occurrences are causally indiscernible. If there is more than one such occurrence, though, then CE is false. For CE claims, contrary to the hypothesis that, say, there is more than one number, that all occurrences with an empty causal profile are identical.

Rather than take this as a refutation of CE, though, I propose that we instead take CE to range over only occurrences that are within the causal nexus, i.e. occurrences that stand in at least one causal relation.

But this raises the question of "causally relevant" occurrences. That is, occurrences that are not themselves of the appropriate ontological category to be causal relata, but which are nonetheless thought to contribute somehow to causal relations. If CE is false for all occurrences that stand in no causal relations, then it is false for occurrences thought to be causally relevant. But many of the discussions presupposing CE concern causally relevant occurrences in addition to or rather than causally related ones.

In a discussion of the problem of qua-causation as it afflicts Donald Davidson's Anomalous Monism, Ernest Sosa raises the issue by analogy[5]:

> I have drawn an analogy between *the relevance of mental properties* to the causal efficacy that AM [Anomalous Monism] grants to mental events and the relevance of loudness to the causal efficacy of a loud shot. Neither the mentality of mental events nor the loudness of the loud shot is causally relevant to the respective relevant effects. ([Sos95]: 41, emphasis added)

The causal relata here are agreed to be events. All the same, the question is posed about the *causal relevance of properties* to the potency of events. CE is assumed insofar as an event's discernible properties are assumed to make discernible causal contributions. This is CE's contrapositive: if two occurrences are discernible, then they are causally discernible. The supposition here, then, is that the discernibility of two occurrences that are not of the category of causal relata entails the discerniblity of their causal relevance.

In discussing the problem of causal exclusion for non-reductive physicalism, Jaegwon Kim also asks after the causal relevance of mental properties

[5]For endorsements of this analogy (or others sufficiently similar) and the moral Sosa draws as a problematic, see [Dre89], [LL87]

as compared to physical properties, even though both accrue to the same causally potent event.

He begins with the problematic.

> . . . the problem of causal exclusion is to answer this question: *Given that every physical event that has a cause has a physical cause, how is a mental cause also possible?* ([Kim98]: 38) (emphasis in the original)

The possibility of a mental cause as well as a physical cause is troubling just in case mental and physical causes are discernible; with this, given CE, the worry is that *distinct* mental and physical causes overdetermine the effect. But non-reductive physicalists do not think that mental and physical causes are discernible. They are physicalists; they claim that the mental causal relata, namely token mental events, are identical to the physical causal relata, token physical events. What, then, is the trouble? Kim answers:

> . . .the question ultimately involves the causal efficacy of mental *properties*, and antireductionism precludes their reductive identification with physical properties. ([Kim98]: 38) (emphasis in the original)

As in Sosa, the discernibility of two non-causal occurrences is alleged to entail the discernibility of their causal relevance.[6]

It is tempting indeed to condemn each and every argument here: *Occurrences that do not stand in causal relations all have the same causal profle—the empty profile. The loudness of Sosa's shot is one such property, the force of the bullet is another. Though these two are causally indiscernible by their blank causal profiles, they are clearly not identical. CE is false.* But, this would preclude a wider and perhaps plausible application of the reasoning behind CE; its application to causal relevance. What we need, then, is a version of CE that applies to occurrences that are not themselves causal relata and, to go with it, a concept of causal indiscernibility for non-causal occurrences.

Let us reserve the phrase "indiscernibility of causal relevance" and its cognates to refer to causal indiscernibility as it applies to non-causal occurrences. I attempt to clarify this concept elsewhere, but let me say something brief about about its relation to the problem of causal exclusion. The principle that parallels CE for non-causal occurrences is as follows; call it the principle of Exhaustion by Causal Relevance:

[6] [Rit05], [Stu05], and [Ben08] make similar assumptions about causal relevance.

ECR If 'a' and 'b' refer to occurrences that are indiscernible in their causal relevance, then a = b.

The problem of causal exclusion is often proposed either as Kim has done above or as a set of inconsistent propositions. Karen Bennett gives a conveniently tidy mock up:

Distinctness Mental properties (and perhaps events) are distinct from physical properties (or events).

Completeness Every physical occurrence has a sufficient physical cause.

Efficacy Mental events sometimes cause physical ones, and sometimes do so in virtue of their mental properties.

Nonoverdetermination The effects of mental causes are not systematically overdetermined; they are not on a par with the deaths of firing squad victims.

Exclusion No effect has more than one sufficient cause unless it is overdetermined. ([Ben08]: 280-1)[7]

Bennett claims that without the **Nonoverdetermination** proposition, there is no problem. Rather, it is simply the case that the effects of mental causes have two sufficient causes. I have objected that there are two causes only if the discernibility of occurrences entails their discernibility as causes. But when it comes to the indiscernibility of causal relevance, there is another pressure point, namely, that occurrences like properties are *not causes*, and they are not supposed to be causes by the philosophers who either pose or fall prey to the problem of causal exclusion.[8] Thus, even if ECR is true and the fact that two properties are discernible entails that they play discernible causal roles, it does not follow that the effects of mental causes have two sufficient *causes*. Rather, there is still only one cause, but this cause has (at least) two causally relevant properties. Where I try to clarify ECR and its attendant concept of indiscernibility, I also try to answer whether this is indeed a case of overdetermination.

[7] See [Rit05] and [Stu05] for similar motivations.

[8] Although things are more complicated than this because many of them subscribe to Kim's view of events, on which a difference in property suffices to bring about a difference in event. Thus, Kim denies what most non-reductive physicalists, i.e. property dualists, affirm—that a single event can have both a mental and physical property. In this case, then, ECR is irrelevant as regards events, but CE comes to bear and ECR may arise for other non-causal entities like facts.

3 Causal Indiscernibility

The overarching concept of causal indiscernibility can be stated simply: *a* and *b* are causally indiscernible only if *a* and *b* are not discerned by the true theory of causation. We must hasten to add, however, that *a* and *b* must both be inside the causal nexus. Beyond this, we must make a number of other clarifications about our analysis.

We shall address here the causal indiscernibility of *individuals*, by which I mean, for example, event-*tokens*, property-*instances*, or particular objects. One may also speak of indiscernible event-*types*, properties, or kinds of objects. The difference between these two, I take it, is simply that the second group can recur in a single world while the first group cannot—they are dated.

Consider event-types and -tokens. The causal profile of an event-type in a world is simply the sum of its tokens' profiles in that world. Thus, two event-types are causally indiscernible just in case their tokens are causally indiscernible. For these purposes, the difference between intra-world recurrences (for types) and inter-world occurrences (for tokens) is immaterial. If tokens are causally indiscernible in their inter-world occurrences, then the tokens of the type will be indiscernible in their intra-world recurrences. For, the circumstances of each recurrence in a world will be duplicated for at least one occurrence of a token in the infinite possible worlds in which non-compossible tokens occur. If an event-type occurs at t_1 and recurs at t_2 in w, there are worlds w_1 and w_2 in which tokens of this type occur in the same circumstances as in w at t_1 and w at t_2—indeed, there are infinitely-many such worlds. Thus, for the sake of expediency, when clarifying causal indiscernibility for a given theory of causation, we shall con centrate on indiscernibility of tokens unless the view is idiosyncratic in its view of type-token relations. We assume that two event-types are indiscernible iff their tokens are indiscernible by necessity in some stable set of possible worlds. So too, *mutatis mutandis*, for properties and their instances, classes of objects and their particulars, etc.

As we saw above, if *a* or *b* is neither a cause nor an effect, then the theory of causation will (presumably) fail to individuate them simply because it does not individuate non-causes/effects. So we cannot say simply that two occurrences are causally indiscernible just in case they are indiscernible from the perspective of the theory of causation. In this case, all (e.g.) properties will be causally indiscernible since they're not causes. This may seem unobjectionable, but it is problematic because it vitiates the intuition that properties can differ in their causal relevance. Although this may be apposite, we should not let our clarification of causal indiscernibility for causes and effects dictate it.

Where *a* and *b* are embedded in the causal nexus, though, another problem arises. We want to remain agnostic about whether *a* and *b* are discernible occurrences even while we affirm that they are causally indiscernible. So if the theory of causation doesn't discern *a* and *b as causes*, we want it to do so without speaking to their discernibility across the board. (If they are discernible as causes, they are of course discernible.) But seldom are theories of causation thus taciturn. Rather, they typically assume that as causal relata are individuated, so are the ontological category of the causal relata and vice versa.

Indeed, one might argue that we discover what is the ontological category of the causal relata partly by considering how the causal relata are individuated. On this view, a theory of how to individuate the ontological category of the causal relata just is a theory of how to individuate the causal relata, and there is no sense to be made of causal indiscernibility independent of the indiscernibility of the ontological category of the causal relata. Consider arguments to the effect that the causal relata are this or that ontological category[9]:

(P1) At least some causal relata are F — $(\exists x)(Rx \,\&\, Fx)$

(P2) Only the elements of ontological category Ω are F — $(x)(Fx \rightarrow Ox)$

(P3) All causal relata are of the same ontological category — $((\exists x)(Rx \,\&\, Ox) \rightarrow (x)(Rx \rightarrow Ox))$

(C1) $\therefore$ At least some causal relata are of ontological category Ω — $(\exists x)(Rx \,\&\, Ox)$ (by P1, P2)

(C2) and $\therefore$ All causal relata are of ontological category Ω — $(x)(Rx \rightarrow Ox)$ (by P3, C1)

It has been argued, for example, that since causal relata seem to be individuated finely, they must be a certain kind of event, namely, the kind that is individuated finely.[10] Fx = x is individuated finely, and Kim events, which are triples of an object, a property, and a time, are taken to fill out the ontological category Ω. Given that all causal relata are from the same ontological category and that some causal relata are Kim events—since they're

[9]See, for example, arguments from absences [Mel95], [Ben88], as well as the "causal differences" argument [Kim93a]

[10] [Kim93a]

individuated finely—it is concluded that all causal relata are Kim events. If this arc of reasoning is our route to discovering the nature of the causal relata, how can we commit to a theory for individuating causal relata without thereby committing to an individuation of the ontological category of the causal relata? CE seems to be true *on methodological grounds*.

It is true that the foregoing argument is valid, and it's true that it is acceptable as a means to discovering the causal relata, but this doesn't suggest that the individuation of the ontological category of the causal relata *just is* the individuation of causal relata. It doesn't suggest that discernibility entails causal discernibility. Rather, it reminds us that causal discernibility entails ontological discernibility, and so causal relata cannot be individuated *more finely* than the ontological category of causal relata. It may nonetheless be the case that occurrences of the ontological category of causal relata are individuated more finely than causal relata.

Consider the argument where Fx = x is individuated coarsely, and which proceeds as follows. It is observed that some causal relata are individuated coarsely, and then proposed that Davidsonian events, which are individuated by their causal profiles ([Dav80b]: 306), are the only occurrences individuated so coarsely of which it makes sense to say that one precedes another. So it seems that some causal relata are Davidson events. Given that all causal relata are of the same ontological category, it follows that all causal relata are Davidson events. But must we now say that all events are individuated so coarsely as Davidson events?

Well, we are not *logically bound* to this position. Our options, regarding Davidson's view in particular, will become clear below, but let us note, in addition, that since it is open to us to individuate causal relata one way while individuating the ontological category of causal relata more finely, we should not want to reject this possibility without reason. In the absence of reasons, then, causal indiscernibility may be defined independently from ontological indiscernibility, even though causally discernible occurrences are thereby ontologically discernible. Thus, although events, say, are the causal relata and events with discernible causal profiles are thereby discernible events, we need not say that all discernible events have discernible causal profiles. (Let Dxy = x is discernible from y, and let Cxy = x is causally discernible from y. The distinction is easily noted: $(x)(Cx \rightarrow Dx)$ is clearly true, but we do not know whether $(x)(Dx \rightarrow Cx)$, even given the aforementioned restrictions on the universe of discourse.)

Coming back to the point, then, we can refine our initial claim: *a* and *b* are causally indiscernible if they are not discerned by the theory of causal *connection*. This is a necessary and sufficient condition, provided that *a* and *b* are within the causal nexus. The details of causal indiscernibility depend, of course, on which is the true theory of causation, but we can nonetheless make a few general remarks about how popular views of the causal connection individuate causal relata independent of their individuation as elements of some ontological category.

3.1 Regularity Theories

In general, where *a* *b*, *c*, and *e* refer to causally embedded occurrences, $\mathrm{T}_c\mathrm{E}$ refers to the complete list of criteria for being a cause of *e* on a theory of the causal connection T, and $\mathrm{T}_e\mathrm{C}$ refers to the complete list of criteria for being an effect of *c* on theory T, *a* and *b* are causally indiscernible iff:

1. For all *e*, *a* satisfies $\mathrm{T}_c\mathrm{E}$ iff *b* satisfies $\mathrm{T}_c\mathrm{E}$, and
2. For all *c*, *a* satisfies $\mathrm{T}_e\mathrm{C}$ iff *b* satisfies $\mathrm{T}_e\mathrm{C}$

The results of this arrangement vary, as one would guess, with theories of the causal connection. It is helpful to observe these variations.

Let us begin, then, with regularity theories of causation. In Hume's famous words:

> ...we can define a cause to be *an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second.* Or in other words *where, if the first object had not been, the second never had existed.* ([Hum99]: 60) (emphasis in original)

As many have observed, there are two definitions here. Let us take them in turn and at face value.

First, Hume gives two conditions for an object c's being a cause of an object e: (i) *e* follows *c*, and (ii) all objects similar to *c* are followed by objects similar to *e*. Let us suppose each of these to be a necessary condition and their conjunction to be sufficient. What is causal indiscernibility on this view? Let us start with what it is for *a* and *b*, where 'a' and 'b' refer to objects, to be indiscernible as regards the cause of *e*. The stated conditions hold for *a* if and only if they hold for *b*. So (1) *e* follows *a* iff it follows *b*, and (2) objects are similar to *a* iff they are similar to *b*, and all objects similar to *a/b* are followed by objects similar to *e*. This calls for discussion.

(1) is problematic. The claim that *e* follows *a and b* brings to mind a temporally extended series of three, with *e* last. There is no constraint on the temporal proximity of *a/b* and *e*, and there is no constraint on intermediaries

between *a/b* and *e*, though there seems to be a need for at least one of these two constraints. In the absence of some such constraint, our account of causal indiscernibility generates a vast number of false positives: the first two links, *x*, *y*, in any three-link causal chain, *x* - *z*, are causally indiscernible as regards the last link, *z*. This lacuna must surely be filled, but the problem is not simply the absence of any *temporal* constraint. For this same liberality is present in Hume's definition. Hume's (i) says only that *e* follows *c*, and this permits the same variety of temporal relations as our version. What we want is that *a* and *b* are indiscernible in their preceding of *e*.

What, then, is it for *a* and *b* to precede *e* indiscernibly? There seems little answer until we recognize that since *a* and *b* are particular objects and Hume has not introduced any modal operators, indiscernibility here is tantamount to observational coincidence in actuality plus the first conjunct of condition (2), that *a* and *b* are similar to all and only the same objects. Let us read observational coincidence as the stronger and more familiar spatiotemporal coincidence. As regards the similarity condition, let us say that *a* and *b* are particulars of the object kinds A and B. And, an object is similar to a particular of A iff it is similar to a particular of B. This shall be true if a_1 is spatiotemporally coincident with b_1, a_2 is spatiotemporally coincident with b_2.... I assume that this is no stronger a claim than that A and B are indiscernible in all nomically possible worlds. So the condition of spatiotemporal coincidence is satisfied as well.

Thus, on this definition of a cause, *a* and *b* are causally indiscernible *as a cause of e* if and only if (i) *a* and *b* are spatiotemporally coincident, and (ii) all and only objects similar to *a* are similar to *b*, and all of these are followed by objects similar to *e*. For *a* and *b* to be causally indiscernible as an effect of *c*, the results are similar: *a* and *b* are spatiotemporally coincident, an object is similar to *a* if and only if it is similar to *b*, and all the objects that precede an object similar to *a* and *b* are similar to *c*. Once again, I propose that spatiotemporal coincidence in all nomically possible worlds suffices to satisfy this condition.

Whether there are any discernible objects that are spatiotemporally coincident is a matter of contention, though it is widely held to be possible.[11] The canonical example is a statue, Goliath, and the lump of clay constituting it, Lumpl. The day before the sculptor forms Lumpl into Goliath, Lumpl exists but Goliath does not. Thus, they are discernible, and from the next day on, they are spatiotemporally coincident.[12] Thus, the causal

[11] See [Gib75], [Rea97]

[12] See [Gib75]

indiscernibility of *a* and *b* does not seem to be decided by the discernibility of *a* and *b*. Goliath and Lumpl are not spatiotemporally coincident *by nomic necessity*, though. They're not even coincident in actuality. The issues are similar in the case of nomically necessary spatiotemporal coincidence, but it is unclear whether they resolve in the same fashion. A very similar issue arises for Davidsonian events, which we shall address shortly, in the context of Davidson's account of causal relations.

3.1.1 Davidsonian Causal Indiscernibility

Before we proceed to Hume's other definition of a cause, let us consider a contemporary attempt to refine Hume's view—Donald Davidson's regularity theory in terms of the "Principle of the Nomological Character of Causality":

PNCC events related as cause and effect fall under strict deterministic laws. ([Dav80c]: 208)

We can understand what it means for an event to "fall under" a law by Davidson's comments about singular causal *statements*:

SCS if 'a caused b' is true, then there are descriptions of *a* and *b* such that the result of substituting them for 'a' and 'b' in 'a caused b' is entailed by true premises of the form of (L) and (P) ([Dav67]: 701)

(L) is a strict deterministic law stating that if an event with certain properties occurs at a time, call it event c^*, it causes an event with certain properties at a later time, call it e^*; (P) is the proposition that c^* occurs at a certain time. ([Dav67]: 699-700) Thus, (L) and (P) entail that c^* caused e^*. An event c is a cause at t, then, just in case the description of the cause in a strict deterministic law (L) is also a description of c, and c occurs at t (P). Similarly, an event e is an effect at $t + n$ just in case it is described by the specification of the effect in a strict deterministic law and it occurs at $t + n$.

Where 'a' and 'b' refer to events, then, *a* and *b* are indiscernible in a case of singular causation as the cause of e at $t + n$ if and only if (i) the description of the cause in a strict deterministic law is a description of both *a* and of *b*, and (ii) both *a* and *b* occur at t.

So far, this is unproblematic, but it threatens to be too liberal as long as the descriptions in strict deterministic laws are left unconstrained. For (ii) demands of *a* and *b* only that they be temporally coincident. Reviewing the details of Davidson's view on strict deterministic laws would take us too far afield, but the upshot is nonetheless clear: all strict laws are either laws of an ideally completed physics or derivable from such laws via bridge laws; the descriptions in strict laws, then, are the descriptions of an ideal physics or descriptions connected to them via bridge laws.[13] Davidson's view on bridge laws, though, is pessimistic—he doubts that bridge laws will extend so far as biological descriptions ([Dav80a]: 241). So let us suppose that only the descriptions of an ideally completed physics appear in strict deterministic laws. Thus, (i) is tantamount to the claim that a and b are indiscernible as the cause of e only if they are both described by the same physical description.

It is, of course, very unclear what follows from fitting the same description from an ideal and as-yet-unknown physics, but let us suppose that an ideal physics will be the fundamental organization of space and time, and that events sharing a description in such a physics will be spatiotemporally coincident. Thus, *a* and *b* are indiscernible as the cause of e if and only if they are spatiotemporally coincident. Further, if *a* and *b* are to be causally indiscernible for some modal strength, then it must be the case in all the worlds possible for that modal strength that (i) *a* is described by a physical description in the cause or effect of a strict law if and only if *b* is, and (ii) *a* occurs at a time *t* if and only if *b* does. Thus, *a* and *b* are nomically causally indiscernible iff they are spatiotemporally coincident in all the nomically possible worlds, they are metaphysically causally indiscernible iff they are spatiotemporally coincident in all the metaphysically possible worlds, and so on. This ensures that *c* is a cause of *a* iff it is a cause of *b*, and that *e* is an effect of *a* iff it is an effect of *b*.

And here arises, as promised, the issue of spatiotemporal coincidence for Davidsonian events. Davidsonian events are individuated by their causal profiles: 'x = y', where 'x' and 'y' refer to events if and only if:

($\mathbf{E}_D$) ((z)(z caused x $\leftrightarrow$ z caused y) & (z)(x caused z $\leftrightarrow$ y caused z)) ([Dav80b]: 306)

This definition ensures that causally indiscernible Davidsonian events are completely indiscernible Davidsonian events, so long as the modal force of each "indiscernible" is the same. For, any Davidsonian events *a* and *b* with the same causal profile will satisfy both conjuncts of the above identity criterion. But suppose, as is plausible for identity claims, E_D operates over all possible worlds. Thus, $a = b$ iff they have the same causal profiles in *all possible worlds*, i.e. just in case they are spatiotemporally coincident in all

[13] See [McL89] 115-6 for an illuminating and brief discussion of Davidson's strict laws; see [McL85]: for an illuminating and extended discussion.

possible worlds. Thus, the complete indiscernibility of Davidsonian events is determined by their causal indiscernibility just in case spatiotemporal coincidence in all possible worlds is necessary for causal indiscernibility.

We shall not address the matter here, but let us adduce a few of the relevant considerations. Prima facie, once we have introduced the apparatus of possible worlds, we should add to Davidson's PNCC that c falls under a law that holds *in the world in which c occurs*. Further, then, only nomically possible worlds are relevant to a Davidsonian event's causal profile—the event isn't a causal relatum where it doesn't fall under a law. But causal indiscernibility may still decide complete indiscernibility. For it may be that (i) the nomically possible worlds are *all the worlds there are*, or it may be that (ii) Davidsonian events exist in only the nomically possible worlds—if, for example, their nomic roles are essential to their identities. In both of these cases, spatiotemporal coincidence in all possible worlds is necessary for the causal indiscernibility of Davidsonian events, but for reasons that require elaboration and justification.

3.2 Lewis's Counterfactual Analysis

We may now turn to Hume's other definition of a cause.

> ...we can define a cause to be ... *where, if the first object had not been, the second never had existed.* ([Hum99]: 60) (emphasis in original)

David Lewis confesses a preference for this definition in his seminal paper, "Causation", and his analysis of it there is, I venture, the best gloss on the view expressed here. Rather than attempt our own analysis of this view, then, let us simply turn to Lewis's.

According to Lewis, if an event e depends causally on an event c in a world w, then c is a cause of e; but if c is a cause of e, it does not follow that e depends causally on c: it is sufficient that there is a causal chain from c to e such that each later link in the chain causally depends on its immediate predecessor. ([Lew73]: 563) Causal dependence is sufficient but not necessary for causation. Still, let us focus simply on Lewis's account of causal dependence; our aim here is thus to clarify the concept of indiscernible causal dependence.

We might then be inclined to define causal indiscernibility in terms of causal dependence, but I prefer to prescind from the difference between these two concepts in our analysis. Indiscernibility with respect to Lewisian causation *prima facie* demands indiscernibility in indefinitely-long chains of causal dependencies; this is unnecessarily restrictive. Instead, we may concern ourselves with the *immediate* causes and effects of occurrences a and b. I believe that the intuitive concepts of immediate cause and immediate effect are approximated by the following proposition: c is an immediate cause of e iff e is causally dependent on c. Strictly speaking, then, we shall be devising a scheme according to which a is indiscernible from b with respect to causal dependencies. Nonetheless, it shall be convenient to refer to indiscernible causal dependence by "causal indiscernibility" and its cognates. What is meant here is that a and b are indiscernible with respect to immediate causes and effects. Let us now turn to our analysis.

An event e depends causally on an event c in a world w if and only if both 1 and 2 hold in w:

1. $(\mathrm{O(c)} \Box \rightarrow \mathrm{O(e)})$
2. $(\neg\mathrm{O(c)} \Box \rightarrow \neg\mathrm{O(e)})$ ([Lew73]: 562-3)

where Ox is the proposition that x occurs. We evaluate 1 and 2 as follows.

First, certain conditions in w make 1 or 2 "automatically true".

(1) $\mathbf{(O(c) \Box \rightarrow O(e))}$ is true if c and e occur in actuality

(2) $\mathbf{(\neg O(c) \Box \rightarrow \neg O(e))}$ is true if neither c nor e occurs in actuality

Either 1 or 2 is "automatically true" for every causally dependent pair of events. Where one is "automatically true", the causal dependence relation holds just in case the other proposition holds. Hence, if c and e occur, e causally depends on c if:

$\mathbf{(\neg O(c) \Box \rightarrow \neg O(e))}$

And if neither c nor e occurs, then e depends causally on c if:

$\mathbf{(O(c) \Box \rightarrow O(e))}$

The remaining counterfactual is then evaluated by reference to the nearest possible world.

$\mathbf{(A \Box \rightarrow C)}$ is true iff

1. There are no possible A-worlds (in which case $(\mathrm{A} \Box \rightarrow \mathrm{C})$ is *vacuous*)
2. some A-world where C holds is closer (to w) than is any A-world where C does not hold

One world is closer to actuality than another if the first is more similar to ours overall than the second. Lewis proposes that we give special weight to "comprehensive and exact similarities of particular fact throughout large spatiotemporal regions." ([Lew73]: 560) Thus, the nearest world where an event holds is a world exactly like ours from the beginning of time up to the time at which, in our world, the event fails to occur. At that instant, instead of failing to occur, the event occurs. So too, *mutatis mutandis*, for the nearest possible world in which an event fails to occur. Thus, Lewis gives us the gloss on counterfactuals:

> In other words, a counterfactual is nonvacuously true iff it takes less of a departure from actuality to make the consequence true along with the antecedent than it does to make the antecedent true without the consequent. ([Lew73]: 560)

3.2.1 Lewisian Causal Indiscernibility

I take it that it does not suffice for the Lewisian causal indiscernibility of a and b that e is simply causally dependent on both a and b in @. Indeed, it should not even suffice that e is causally dependent on a and b in @ and a and b are temporally coincident. Neither of these conditions distinguishes causally indiscernible occurrences from joint causes or overdeterminative causes. We revisit the relevant distinctions in section 4, but for now let us just add that causally indiscernible occurrences must be causes *for the same reasons*. That is to say, since it must be true that e is causally dependent on a iff e is causally dependent on b, it must be that the truth-makers for the counterfactuals for e's being causally dependent on a also make true the counterfactuals for e's being causally dependent on b.

For singular causation, then, where 'a' and 'b' refer to events, e's causal dependence is indiscernible between a and b in a world w if and only if 1 and 2 hold for both a & e in w and b & e in w. We may thus specify 1_{CD} and 2_{CD} for a world w.

$\mathbf{1}_{CD}$ $(O(a) \Box\!\!\rightarrow O(e))$ & $(O(b) \Box\!\!\rightarrow O(e))$

$\mathbf{2}_{CD}$ $(\neg O(a) \Box\!\!\rightarrow \neg O(e))$ & $(\neg O(b) \Box\!\!\rightarrow \neg O(e))$

Once again, there are two cases. The case in which 1_{CD} is automatically true and that in which 2_{CD} is automatically true. Where 1_{CD} is automatically true, O(a), O(b), and O(e) are all true in w; thus, a and b are causally indiscernible in w if the world nearest to w in which a fails to occur is also the nearest world in which b fails to occur, and e fails to occur there as well.

Where 2_{CD} is automatically true in w, all three events are absent from w, and a and b are causally indiscernible if the world nearest to w in which a occurs is also the nearest world in which b occurs, and e occurs there as well.

Focusing on the relation between *a* and *b*, we may say, then, that a is causally indiscernible from b in the actual world, @, *only if a and b are coincident* in some set of worlds that includes (i) @ and (ii) w*, the world nearest to @ in which (O(a) ∨ O(b)) has the opposite truth-value it does in @. *a* and *b* must be coincident in @ if either 1_{CD} or 2_{CD} is to be automatically true in actuality; they must be coincident in w* if whichever of 1_{CD} and 2_{CD} is not *automatically* true in actuality is to be true in @ nonetheless.

For example, if O(a) O(b), and O(e) are all true in @, in accordance with (i), then 1_{CD} is automatically true in @. If 2_{CD} is to be true in @, then it must be that the nearest world in which *either* ¬O(a) or ¬O(b) is true is a world were *both* ¬O(a) and ¬O(b) are true. If the world nearest to @ in which ¬O(a) ∨ ¬O(b) is one in which, say, ((¬O(a) & ¬O(e)) & O(b)) is true, then the first conjunct of 1_{CD} shall be true in @ but the second is false. In accordance with (ii), then, a and b must be coincident in the nearest world w* where O(a) ∨ O(b) is true.

The coincidence of *a* and *b* in @ and w* entails that (iii) a and b are coincident in all the worlds between @ and w*. If they are not coincident in one of these intermediate worlds, then one occurs and the other does not. Thus, O(a) ∨ O(b) shall be true, but a and b will not be coincident in the nearest world where this is so, contra (ii).

Call this set of worlds in which *a* and *b* must be coincident to be causally indiscernible in @ the *Causal Set* for *a* and *b* in @. If *a* and *b* are coincident in the Causal Set for *a* and *b* in @, and if 1_{CD} and 2_{CD} are true, then *e*'s causal dependence is indiscernible between *a* and *b* in @. Moreover, it shall be the case that for *any e*, *e* is an effect of *a* iff it is an effect of *b*. This does not yet suffice for the causal indiscernibility of *a* and *b* in @, though. They must also be causally dependent on the same events; we might say that they must be effects of the same cause(s).

This does not mean that the cause of *a* must *be identical to* the cause of *b*; rather, it means only that the cause of *a* must be causally indiscernible from the cause of *b*. If CE is true, then these claims are equivalent, and it is immaterial for all but rhetorical purposes how we specify the desideratum. If CE is false, then it is logically possible that there are discernibles that are causally indiscernible. The discernibility of any such occurrences, however, shall hold by virtue of facts that are irrelevant to their causal roles—otherwise they would not be causally indiscernible. Since these facts do not affect the causal profiles of the discernibles in question, I see no reason to think that they should influence the causal profiles of the *effects* of such causally indiscernible discernibles. In the absence of such reasons, I propose that *a* and *b* may be indiscernible with respect to the event(s) on which they causally depend only if the event(s) on which they depend are themselves causally indiscernible. We know what this requires; so let us turn to the requirements peculiar to *a* and *b* as indiscernible effects.

It must be, of course, that *a* is causally dependent on *c* iff *b* is causally dependent on *c*. In terms of possible worlds: for any *c* that is coincident with *a* and *b* in actuality, *a* and *b* are coincident in the nearest world where O(c) has the opposite truth-value it does in actuality. Suppose O(c), O(a), and O(b) are all true in @, and let w* be the nearest world in which ¬O(c) is true. Whether *a* and *b* causally depend on *c* is immaterial as regards causal indiscernibility; what's important is that *a* does iff *b* does. So it matters little whether O(a) & O(b) is true in w*, as long as it is true that O(a) ↔ O(b) in w*. The results are the same if it is true in actuality that ¬O(a) & ¬O(b) & ¬O(c): in the nearest world in which O(c) is true, O(a) ↔ O(b) is true.

Considering the breadth of events that are coincident in actuality with any given event, this constraint bears a redoubtable heft. Setting aside for now the truth-values of O(a) and O(b) in the nearby worlds, suppose that the world w* for any given event *c* is the world that differs from @ with respect to *c only*. If this is true, then the set of worlds that determine the causal dependencies in @ is the set of {@} ∪ {the worlds that differ from @ with respect to a single event *c* alone}. Call this set of worlds the *Effectual Set* for *a* and *b* in @. If O(a) & O(b) is true in @, then for every *c* of which O(c) is true in @, there is a world in the Effectual Set for *a* and *b* in @ that is as similar to @ as possible except that ¬O(c) is true. And if ¬O(a) & ¬O(b) is true in @, then for every *c* of which ¬O(c) is true in @, there is a world in the Effectual set for *a* and *b* in @ that departs from @ as little as possible while O(c) is true.

Since inclusion is determined by the presence or absence of single event-tokens in a world, the extent of the Effectual Set for events in @ depends on how events are individuated. It is of particular importance whether events may be logically compound. If there are events *c & d*, *c* ∨ *d*, etc. where *c* and *d* are also events, then the Effectual Sets for events in @ will include much more than what we might consider a thin shell of worlds around @. Suppose, for example, that *a* and *b* occur in @, then for some actual *c*, the nearest world without *c*, w*, is in the Effectual Set for *a* and *b* in @. If there are logically compound events, then the second-nearest world without *c* may also find its way into the Effectual Set. Call this world w′. As we did for w*, let us assume that w′ differs from w* in only one respect: O(d) has the opposite truth-value there. Suppose this means that ¬O(d) is true in w′. Presumably, then, w′ is the world nearest to @ in which ¬O(c ∨ d) is true. And since O(c ∨ d) is true in @, the event *c* ∨ *d* is coincident with *a* and *b* in @. w′ is thus included in the Effectual Set for *a* and *b* in @ because it is the world that determines whether *a/b* causally depends on *c* ∨ *d*.

The method used to bring w′ into the Effectual Set for *a* and *b* in @ can of course be iterated to bring the web of worlds branching out from w* into the set as well. Similarly, all other worlds may be brought into the Effectual Set for *a* and *b* in @ by disjoining, conjoining, and negating individual events.

Suppose again that *a*, *b*, and *c* occur in @, suppose that *d* does not occur in @, and let w# be the nearest world in which *d* occurs. Since *d* is not coincident with *a* and *b* in @, it seems that w# is not in the Effectual Set for *a* and *b* in @. Since the counterfactual ¬O(d) □→ ¬O(a/b) is false in @, *d* is ineligible to be a cause of *a/b*, and so it seems irrelevant what happens in the nearest world in which O(d) is true. Whether O(d) □→ O(a/b) is true there or not, since ¬O(d) □→ ¬O(a/b) is false in @, *a* and *b* cannot be causally dependent on *d* in @. Moreover, since w# is the nearest world in which *d* occurs, it presumably differs from @ in this respect only. If this is so, then all the events that *are coincident with a and b in @* also occur in w#, and thus w# does not seem to determine the truth of *any* of the counterfactuals that determine causal dependencies in @, and so it seems w# cannot decide the causal indiscernibility of *a* and *b*.

Or so it may seem. If we may conjoin and negate events, then *c & ¬d* and *c* ∨ ¬*d* are events. Indeed, they are both events that are coincident with *a* and *b* in @. Thus, both of the following propositions are automatically true in @.

1. O(c & ¬d) □→ O(a/b), and
2. O(c) ∨ ¬d □→ O(a/b)

a and *b* are thus causally dependent on *c & ¬d* if the following is true in @:

(c & ¬d) ¬O(c & ¬d) □→ ¬O(a/b)

And, *a* and *b* are causally dependent on *c* ∨ ¬*d* if this proposition is true in @:

(c *vee* ¬d) ¬O(c ∨ ¬d) □→ ¬O(a/b) is true in @.

Of course, these latter propositions are true in @ iff their respective consequents are true in the nearest world in which the respective antecedents are true: ¬O(a/b) ¬O(c & ¬d) is true in the nearest world in which ¬O(c & ¬d) is true or in the nearest world in which ¬O(c ∨ ¬d) is true. Since the causal dependencies of *a* and *b* are partly determined by these worlds, they must be included in the Effectual Set for *a* and *b* in @.

And one of these two worlds is w#. If it takes less of a departure from actuality to make O(d) true while O(c) is also true, then then w# is the nearest world in which ¬O(c & ¬d) is true, and so it is included in the Effectual Set for *a* and *b* in @. If it takes less of a departure from actuality to make O(d) true while O(c) is false, then w# is the nearest world in which ¬O(c ∨ ¬d) is true, and w# is once again included in the Effectual Set for *a* and *b* in @. Thus, w# is included in the Effectual Set for *a* and *b* in @.

Taken together, the two foregoing points reveal that an event *d* need not be coincident with *a* and *b* in order for the nearest world in which O(d) has the opposite truth-value it does in @, w*, to be included in the Effectual Set for *a* and *b* in @. Rather, for any event *d* that is not coincident with *a* and *b* in @ and any event *c* that is coincident with *a* and *b* in @, there are a number of logically compound events connecting *c* and *d* that are coincident with *a* and *b*, and for some of these compound events, the nearest world in which they occur or fail to occur is also the nearest world in which *d* occurs or fails to occur.

Suppose a, b, and c all occur in @ and d does not. In this case, the compound events c & ¬d and c ∨ d also occur in @, and so these two events are coincident with a and b. Let w* be the nearest world in which d occurs. Now, either c occurs in w* or it does not. If it does, then w* is included in the Effectual Set for a and b in @ because it is the nearest world in which c & ¬d fails to occur. If c does not occur in w*, then w* is included because it is the nearest world in which c ∨ d fails to occur.

Suppose on the other hand that a, b, and c are all absent from @ while d is present. In this case, c & ¬d is also absent from @ and so is c ∨ ¬d. Let w* be the nearest world in which d is absent. Since c & ¬d and c ∨ ¬d are absent from @, they are coincident with a and b in @, and so the following two propositions are automatically true.

1. ¬O(c & ¬d) □ → ¬O(a/b)
2. ¬O(c ∨ ¬d) □ → ¬O(a/b)

a and b are causally indiscernible in @, then, only if a and b are coincident in the nearest world where the antecedent of 1 is false and in the nearest world where the antecedent of 2 is false. One of these two worlds is w*: either c occurs in w* or it does not. If it does not, then w* is the world nearest to @ in which the antecedent of 2 is false, i.e. w* is the world where O(c ∨ ¬d) is true. For if w* is nearer to actuality than any world in which c occurs, then the nearest world in which O(c ∨ ¬d) is true is simply the nearest world in which d and c are both absent. This, of course, is also the nearest world in which O(c ∨ ¬d) is true by dint of the truth of ¬O(d). If c does occur in w*, then w* is included in the Effectual Set for a and b in @ because it is the nearest world in which c & ¬d is present. For, if it takes less of a departure from reality to make ¬O(d) true along with O(c) than to make ¬O(d) true alone, then the nearest world in which ¬O(d) is true, w*, shall also be the nearest world in which O(c & ¬d) is true. And this, of course, is simply the nearest world in which the antecedent of 1 is false, so this world is included in the Effectual Set for a and b in @.

We have just shown that for any event c, the world nearest to @ in which the truth-value of O(c) changes is included in the Effectual Set for a and b in @. This permits us to show that *every* world is included in the Effectual Set for a and b in @ if there are logically compound events. First, let us assume that for every world w′ that is not identical to @, there is at least one event c that either occurs in @ and fails to occur in w′ or fails to occur in @ and occurs in w′. And let us re-introduce our earlier assumption that a world w* that differs from actuality in the presence or absence of a single event c alone is the world nearest to actuality in which the truth-value of O(c) changes its value from @. Given these two assumptions, then, in order to show that every world is in the Effectual Set for a and b in @, we need show only that every world differs from actuality in the presence or absence of a single event c. If this is true, then it follows that every world is the nearest to actuality in which the truth-value of O(c) changes from its value in @, and since we have already shown that every such world is within the Effectual Set for a and b in @, we will have thus shown every world is in the Effectual Set for a and b in @.

With the logical operators at our disposal, there is little difficulty in showing that every world differs from actuality in the presence or absence of a single event c. For any world w′ that differs from @ in at least one event c that either occurs in @ and fails to occur in w′ or fails to occur in @ and occurs in w′, we list the events c, d, e, . . . that occur in w′ but not @. This list may be empty. Next, we list the events f, g, h, . . . that fail to occur in w′ but occur in @. This list may also be empty. Now, there is an event that is the conjunction of all the events in the first list and the complements of all the events in the second list: (c & d & e & ... & ¬f & ¬g & ¬h & ...). w′ differs from @ in this event alone. It occurs in w′ and not in @. Thus, w′ is the nearest world in which this event occurs, and thus, by our earlier observations, it is in the Effectual Set for a and b in @.

If logically compound events are permitted, then, and if there are no constraints on how we might compound events, then it seems that a and b are causally indiscernible—even weaker: they are indiscernible in their causal dependencies as effects—only if they are coincident in all possible worlds. If no such logical compounding of events is permitted, however, then the requirements on the causal indiscernibility of a and b are much weaker.

Consider the Effectual Set for a and b in @ where it is true in @ that O(a) & O(b). The Effectual Set includes only worlds in which ¬O(c) is true for some c of which O(c) is true in @. With the exception of @, the facts in worlds in which O(c) is true are irrelevant, and thus it is immaterial whether a and b are coincident in them. If we can assume that some of these worlds are nomically possible—the world, for example, that differs from @ only in having one additional water molecule—then it follows that a and b need not be coincident by nomic necessity. Similarly, mutatis mutandis, if ¬O(a) & ¬O(b) is true in @.

Finally, let us say that the union of the Causal Set for a and b in @ and the Effectual Set for a and b in @ gives us the *Causal Indiscernibility Set* for a and b in @. Where 'a' and 'b' refer to events, then, a and b are causally indiscernible in @ iff:

LCI$_{@}$ a and b are coincident in the Causal Inscernibility Set for a and b in @

Prima facie, there is no entailment from the Lewisian causal indiscernibility of a and b in @ to the complete indiscernibility of a and b. I assume that indiscernibles are spatiotemporally coincident by logical necessity. It is thus inconsistent with the indiscernibility of *a* and *b* that *a* and *b* fail to be spatiotemporally coincident. Prima facie, however, it is consistent with the Lewisian causal indiscernibility of a and b in @ that a and b fail to be spatiotemporally coincident *even in @*. Even if logically compound events are permitted, all the coincidence that the causal indiscernibility of a and b in @ requires of a and b is worldly coincidence in the Causal Indiscernibility Set. a and b may be coincident throughout all these worlds even while occurring at opposite ends of some or all of them.

Moreover, the logical possibility of a world w# in which O(a) & ¬O(b) is true is inconsistent with the indiscernibility of a and b. If logically compound events are outlawed, though, then w# is not inconsistent with the causal indiscernibility of a and b in @ so long as w# is not in the Causal Indiscernibility Set for a and b in @. Where a and b are causally indiscernible in @ and there is a logically possible world w# outside the Causal Indiscernibility Set, then, a and b shall be causally indiscernible but discernible.

If there are good reasons to think that the Lewisian causal indiscernibility of a and b entails the complete indiscernibility of a and b, they are unknown to me. Making the case, though, would require proving a number of points. Let us array them.

First, it would have to be shown that the causal indiscernibility of a and b demands the spatiotemporal coincidence of a and b. Perhaps there is a similarity ranking for possible worlds that can make this demand on causal indiscernibles. One strategy is to emphasize what Lewis did: long stretches of spatio-temporal similarity. Where such stretches are to be preserved, the worlds nearest to one another will diverge as little as possible in such stretches. Where a and b are spatiotemporally coincident, then, a change in the truth value of (O(a/b)) from @ will depart less from actuality when a and b are spatiotemporally coincident. For there will be a change in one space-time region rather than two. But as far as I can tell, this reason does not go far enough. It suggests that *many* Lewisian causal indiscernibles (supposing there are some) will be spatio-temporally coincident, but it doesn't require it. Require it it must, however, if the relation is to entail complete indiscernibility.

Second, it must be shown that the Causal Indiscernibility Set extends to all logically possible worlds. As I said above, indiscernibles are spatiotemporally coincident by logical necessity. In the idiom of possible worlds, this means that they are spatio-temporally coincident in all worlds consistent with the laws of logic. Causal indiscernibles, however, need be spatiotemporally coincident in only the worlds of the Causal Indiscernibility Set. So the two sets must be shown to include all and only the same possible worlds. As we know, permitting certain logically compound events to be causal relata will show exactly this.

Without committing to logically compound events, however, one might pursue this point with a two-pronged attack. The first prong is to "expand" the Causal Indiscernibility Set out to, say, all the nomically possible worlds. One might try to justify this by claiming that we're interested, for one reason or another, not just in causally indiscernible tokens, but in event-*types* that are causally indiscernible. And it may then be added that since the types subsuming causally potent tokens are tokened in a number of nomically possible situations the Causal Indiscernibility Set for types must be the union of the Causal Indiscernibility Sets for a and b in all worlds in which a token of the type subsuming a/b is a Lewisian cause or effect, i.e. all the worlds in which 1_{CD} and 2_{CD} are true for any tokens of the types subsuming a/b. Since this may include worlds in which *a* and *b* occur *as well as* those in which *a* and *b* do not, it may overcome the limitations mentioned above.

In addition, one may try to “shrink” the set of possible worlds in which a and b occur to fit the set of nomically possible worlds. There are at least two strategies for showing this. On the one hand, one may argue that there are no nomically impossible worlds—the actual laws hold with metaphysical necessity, and metaphysical necessity is equivalent to the strongest necessity there is. On the other hand, one might argue that the causal relations or causal properties of causally embedded occurrences are essential to those occurrences. In any worlds where it is not possible to have these causal properties or relations, then, the relevant occurrences shall not occur. Conjoining either of these strategies with (i) the “expanding” of the Causal Indiscernibility Set and (ii) any similarity relation that would require causal indiscernibles to be spatiotemporally coincident in the Causal Indiscernibility Set would suffice to require that causal indiscernibles are spatiotemporally coincident in all possible worlds.

This bring us to the third point that must be made: additional reasons must be adduced to overcome the claim that necessary spatiotemporal coincidence does not entail complete indiscernibility. This may be approached through a general strategy that licenses the inference from necessary spatiotemporal coincidence to complete indiscernibility or via a particular strategy that licenses the relevant inference for whatever occurrences are at issue. I have no good ideas as to how one might pursue the general strategy. One may attack the particular strategy, however, by arguing that the occurrences in question are individuated by their spatiotemporal positions or some criterion equally restrictive or more so.

4 Acting as a Single Cause

We have specified the relation between causal indiscernibles. We should now articulate a more general relation between causal indiscernibles and their effects. In particular, we should distinguish the causal indiscernibles-effect relation from (1) the joint cause-effect relation, (2) the overdetermining causes-effect relation, and (3) the cause/epiphenomenon-effect relation.

In order to discuss these issues more easily, let us introduce the notion of occurrences "acting as a single cause". If a and b are causally indiscernible (in actuality), then a and b act as a single cause (in actuality). Being causally indiscernible is a sufficient condition for acting as a single cause. I am agnostic as to whether it is a necessary condition. There may be other conditions that suffice for *a* and *b* to act as a single cause, but I am unaware of any.

The primary definition of causal indiscernibility is for singular occurrences. Where singular occurrences are event-tokens, and where all tokens a and b of types A and B are causally indiscernible, though, we may say that the types A and B are causally indiscernible. In this case, the causally indiscernible tokens of A and B, a_1/b_1 a_2/b_2, etc., act as a single cause. Similarly, *mutatis mutandis*, for any causal relata that mimic the token-type relation.

Let us now distinguish the case in which a and b act as a single cause from those cases in which (1) u and v jointly cause an effect, (2) x and y overdetermine an effect, and (3) r causes an effect and s is an epiphenomenon of r.

4.1 Joint Causes, Single Causes

Consider the case in which u and v jointly cause an effect. Let us define the 2^{nd}-order event c as follows: $\Box(c \leftrightarrow (u \,\&\, v))$, where '$\Box$' denotes logical necessity. That is, c is the event of both u and v occurring in the same world.[14] And let c be a cause of e in a world w: for whatever theory of causation is true, let c and e satisfy the criteria for c's being a cause of e in w on this theory.

In this case, let us say that u and v jointly cause e in w iff (i) both u and v are logically necessary but insufficient for the occurrence of c, as follows from the definition of c, and (ii) it is not guaranteed by the force of necessity appropriate to the true causal theory either that u is sufficient for v or that v is sufficient for u.

Let us be more explicit about these conditions. u (v) is necessary for

[14]It is tempting but over-restrictive to add that c occurs at t just in case u and v occur at t. It is tempting partly because on the definition offered, c may simply be a necessary effect of u, which is itself a necessary effect of v. This qualification is over-restrictive, however, because we should not rule out in advance that earlier links in a causal chain are joint causes of later links. To shirk the issue, I have proposed that $c \leftrightarrow (u \,\&\, v)$ holds with logical necessity. Intuitively (but perhaps only intuitively), logical necessitation is stronger than causal necessitation; if this is true, then the relation between c and (u & v) cannot be one of earlier and later links in a causal chain—such a relation would be too weak.

the occurrence of c iff: in all the logically possible worlds, if u (v) does not occur, then c does not occur. u (v) is insufficient for c iff: there are logically possible worlds in which u (v) occurs but c does not.

Condition (ii) says that there are possible worlds (1) in which u occurs without v and (2) in which v occurs without u, and these possible worlds are members of the set of possible worlds which make it true that c is a cause of e. If Davidson's theory of causation is true, then this means that these worlds are nomically possible. If Lewis's theory is true, then these worlds are in the Causal Set for c in w. Call this modal strength that is calibrated to theories of causation *Causal*; here, it is causally possible and not causally necessary that u occurs without v and that v occurs without u.

Consider the causally possible world in which u occurs without v; call this world w*. c does not occur in w* because v is absent, and v is logically necessary for c. Similarly, *mutatis mutandis*, for the world in which v occurs without u. Thus, it follows from (i) and (ii) that it is a matter of not only logical but also causal possibility that u and v occur without one another and thus without c.

I take it that this is tantamount to saying that neither u nor v is itself a cause of e in w. And this is how it should be: joint causes are not causes except ... *jointly*. Where a and b are causally indiscernible, however, *both* a and b are causes of e in w. This is not to say that a or b is sufficient to bring about e on its own—it may be that causes are not sufficient for their effects. Rather, as we have seen, causal indiscernibles both satisfy the criteria for being a cause. Using e for the effect in a world w, we would say: for whatever theory of causation is true, a and b both satisfy the criteria for being a cause of e in w on this theory.

4.2 Overdetermination, Single Causes

If both a and b are sufficient causes of e, however, then e seems to be overdetermined. There is extensive debate over the conditions that are sufficient for overdetermination. Since we will attempt to show that a and b do not overdetermine an effect e, we shall be concerned only with the necessary conditions for overdetermination. In particular, we aim to show that a and b do not satisfy a necessary condition for overdetermination.

Let me propose that where x and y overdetermine an effect e in w, it is necessary that either x or y is *causally unnecessary* for e in w. As stated, this is ambiguous for Lewisian causation. For Davidsonian causation, it is not: x and y overdetermine e only if there is a nomically possible world in which either e occurs without y or in which e occurs without x. As regards

Lewisian causation, however, we have not yet specified the Causal Set—is it the Causal Set for x in w or for y in w?

What I have in mind is that e occur either without x or without y in the *union* of the Causal Set for x in w and the Causal Set for y in w. The effect of this constraint is to ensure that it is causally possible that at least one of the causes, x or y, is a sufficient cause of e *in the absence of the other cause.* The same goes for Davidsonian causation: it must be nomically/causally possible that at least one of the causes brings about the effect in the absence of the other.

The idea here is simply this. An effect is overdetermined if at least one of its causes *would have, in the absence of the other*, sufficed to bring about the effect. We render "would have" in terms of counterfactual possibilities, namely, the cases in which one of the causes occurs without the other. And, further, these possibilities are the very ones that bear on the causal relations in w—i.e. they are the possibilities that fill out *causal possibility.*

If a and b are causally indiscernible as the cause of e in w, however, then a $\leftrightarrow$ b holds with causal necessity, and so there is no world in the Causal Set for a and b in w or in the set of nomically possible worlds in which either (i) e occurs with a and not b or (ii) e occurs with b and not a. For a and b always co-occur in these worlds.

Compare the necessary condition on overdetermination proposed by Karen Bennett:

> e is overdetermined by c_1 and c_2 only if:
>
> **(O1)** if c_1 were to occur without c_2, e would still have occurred: $(c_1 \,\&\, \neg c_2) \,\Box\!\rightarrow e$, *and*
>
> **(O2)** if c_2 were to occur without c_1, e would still have occurred: $(c_2 \,\&\, \neg c_1) \,\Box\!\rightarrow e$ ([Ben03]: 476)

Bennett justifies O1 and O2 as necessary conditions by the following reasoning:

> The main reason is simply that they capture the reasoning we engage in when we want to distinguish cases of genuine overdetermination from cases of joint causation, or from cases in which one of the putative causes is not really a cause at all. Let c_1 and c_2 be the shots fired by two members of a firing squad, and e be the victim's death. If we needed to decide whether or not the death was overdetermined, we would ask precisely whether these two counterfactuals are true...If the answer to both questions is

'no'—if both counterfactuals are false—then the death was not overdetermined, for it was jointly caused by the two gunshots. If only one of the counterfactuals is false, at most only one of the gunmen is guilty...It is hard to see how [these counterfactuals] could *fail* to be true when the relevant e is overdetermined. ([Ben03]: 477)

Finally, Bennett adds that not only must O1 and O2 be true, both must be true *nonvacuously* in w. If there is no causally possible world in which, say, the antecedent of O1 is satisfied, then O1 is true only vacuously, and thus c_1 and c_2 do not overdetermine e. This is the case of course if, say, c_1 cannot occur without c_2 (if $c_1 \rightarrow c_2$) or the reverse (if $c_2 \rightarrow c_1$).

Thus, Bennett's necessary condition fails to hold iff: (i) either c_1 or c_2 is insufficient for e, or (ii) either $c_1 \rightarrow c_2$ or $c_2 \rightarrow c_1$. Notice that this necessary condition is stronger than the condition proposed above. There, *only one* of O1 and O2 must be true nonvacuously. Since causal indiscernibles fail to meet the weaker condition, they fail to meet Bennett's stronger condition as well.

If c_1 and c_2 are causal indiscernibles in w, then $c_2 \leftrightarrow c_1$ holds in all causally possible worlds. It is thus causally impossible to satisfy the antecedent of either O1 or O2, and they are both vacuously true in w. This is unsurprising if CE is true—a and b do not overdetermine an effect when a = b. But once again, the interesting case is whether it is possible that a $\neq$ b. If this is possible, then it is an interesting result that discernible occurrences do not overdetermine an effect. We may put a finer point on why this may be, if it is, in the next section.

4.3 Single Causes

Sailing fairly cleanly between the Scylla and Charybdis of joint causes and overdetermining causes, then, we may characterize causal indiscernibles and their effects as follows. Once again, let us define a 2^{nd}-order event c: $\Box$(c $\leftrightarrow$ (a & b)); but here, let '$\Box$' denote *causal* necessity. So, c is the event of both a and b occurring in the same world. And let c again be a cause of e in a world w: for whatever theory of causation is true, let c and e satisfy the criteria for c's being a cause of e in w on this theory.

In this case, let us say that a and b act as a single cause of e in w iff (i) both a and b are sufficient and causally necessary for the occurrence of c, and (ii) it is causally necessary in w that a $\leftrightarrow$ b.

Once again, a (or b) is causally necessary for the occurrence of c iff: in all the causally possible worlds for c in w, if a (or b) does not occur, then c does not occur. And a (or b) is sufficient for c iff: there are no causallypossible worlds for c in w in which a (or b) occurs but c does not.

Condition (ii) says that there are no causally possible worlds for c in w in which a occurs without b; this we already know. Still, it serves here to explain the startling relationship between a, b, and c in the causal possibilities surrounding w. It may be mysterious how a could be sufficient for c if b is necessary; surely a alone cannot suffice for c if c cannot occur without b. a does not need to suffice for c *alone*, however, in order to suffice for c. Rather, it need only be that a → c holds by causal necessity, as it does. Similarly, it may be puzzling how b can be necessary for c if a is sufficient for c, but b need not prevent any other occurrences from sufficing for c in order to be necessary for c. It need be only that c → b holds by causal necessity.

The apparatus of possible worlds is powerful and useful partly because it reveals and clarifies such distinctions. Indeed, these distinctions have permitted us to discover an arrangement between cause and effect that has hitherto been obscured. Occurrences that jointly cause an effect are both necessary but insufficient for the presence of a cause. Occurrences that overdetermine an effect are both sufficient but unnecessary for the presence of a cause. And, finally, occurrences that act as a single cause are each both necessary and sufficient for the presence of a cause *in the causally possible worlds.*

The following is a list of the salient relations that hold between *a*, *b*, and *c by causal necessity.* The contrast with what holds between them by *logical* necessity should be apparent.

1. a ↔ c
2. b ↔ c
3. a ↔ b
4. (a ∨ b) ↔ c

4.4 Single Causes and Epiphenomena

Using the same conceptual apparatus that has served us thus far, let us now contrast single causes from cases in which a cause, r, and an epiphenomenon, s, co-occur preceding an effect, e. Our strategy here is similar to that pursued in the section on overdetermination. We propose a necessary condition on epiphenomena, and we show that causal indiscernibles fail to meet this criterion.

Let us begin with a plausible and strong claim about epiphenomena: s is an epiphenomenon in a world w only if there is no e such that s is a cause of e in w. This is the general definition of epiphenomena—they are causally impotent occurrences. What is at issue in cases of causal indiscernibles, however, is likely more specific. It must be shown that each of the causal indiscernibles is a cause of *the same effect*, not that each is the cause of *some* effect. This permits us to focus: s is an epiphenomenon *with respect to e in w* only if, for whatever theory of causation is true, s and e fail to satisfy the criteria for s's being a cause of e in w on this theory.

In terms of the events that we have referenced repeatedly in this section, s and a cause r would bear the following relations to the 2^{nd} order event c, which is a cause of e, in the causally possible worlds that legitimate c's status as a cause in w:

1. r suffices for c
2. s does not suffice for c

r suffices for c iff 'r $\rightarrow$ c' holds by causal necessity. s fails to suffice for c iff there is a causally possible world in which s occurs but c does not. Since c is simply a cause of e in w, I take this to be tantamount to saying that s is not a cause of e in w. That is, s is an epiphenomenon.

Notice that this arrangement holds for the paradigmatic examples of epiphenomena. The steam from a locomotive is not causally impotent with respect to all effects, of course, merely with respect to the motion of the train. And, with respect to the motion of the train, there is a causally possible world in which the steam is present but the movement of the train is not. Similarly, *mutatis mutandis* for the movements of a shadow with respect the object casting it.

Causal indiscernibles a and b in a world w, on the other hand, both satisfy the criteria for being a cause of e in a world w. It is causally impossible for either a or b to occur without e. This suffices to distinguish the causal indiscernibles-effect relation from the cause/epiphenomenon-effect relation. To illustrate the point, notice that the disjunction of a pair of causal indiscernibles a and b suffices for c but the disjunction of a cause & epiphenomenon pair s and r does not.

1. $(a \vee b) \rightarrow c$
2. $\neg((s \vee r) \rightarrow c)$

4.5 Single Causes and "Nomic Equivalents"

The preceding has distinguished occurrences that act as a single cause from those that act as a joint cause, those that overdetermine their effects, and those related as cause to epiphenomenon. In addition, it has made explicit the relation between such single cause occurrences in terms of causal necessity.

Nonetheless, I fear that this will not satisfy many of those who worry about the relations between occurrences like a and b. One might ask: given that the occurrence of a suffices to bring about e, what "causal work" is there for b to do? The implication is something like a dilemma. Given that there is no "causal work" for b to do, b must either do causal work already being done, in which case e is overdetermined, or it must do no work, in which case b is epiphenomenal. Finally, if we deny both horns of the dilemma, one is tempted to say that a "didn't really" suffice for e in the first place; hence, a and b jointly cause e.

To stick to the metaphor, the error is to assume that every causal job is single occupancy. This may be the case, but it isn't obvious. If two occurrences may work the same job, the dilemma is only apparent. Since only one cause is present, the effect is not overdetermined. Since each occurrence is a cause, neither is epiphenomenal. And since both a and b are sufficient to bring about e, it is not the case that they are joint causes.

It is illuminating—and I hope ultimately compelling—to compare the remarks in [Kim89]: 86. Kim is there addressing a situation proposed in [Gol69] in which two occurrences are alleged to be "simultaneous nomic equivalents" such that each is a "complete and independent" cause of some effect. Kim claims that this situation is "unstable", and the "instability can be seen in various ways." He then enumerates those ways, of which there are two. First, Kim asks why the effect in the situation is not overdetermined. Second, he asks why the causes do not form a single, jointly sufficient cause.

The foregoing is, in effect, an answer to Kim's queries as they might apply to causally indiscernible occurrences, which act as a single cause, rather than nomic equivalents, which are alleged to be discernible causes.

Kim himself claims that:

> ...these perplexities are removed only when we have an account of the relation between...the two supposed causes of a single action...[and Kim shall argue that] an account that is adequate to this task will show that [the two supposed causes] could not each constitute a *complete* and *independent* explanation of the [effect]. ([Kim89]: 87; *emphasis* in original).

First, let us read "...complete and independent explanation of the effect" as "...sufficient and discernible cause of the effect." Kim discusses explanations and causes together throughout the article, and he often refers to one or the other explicitly while implicitly referring to both. Indeed, he has here rendered Goldman's remarks on causes as pertaining to explanation. So I do not think it is infelicitous to swap "cause" for "explanation" and "sufficient" for "complete". I propose that we also switch "independent" for "discernible" simply in order to render the discussion tractable.[15]

We have given an account of the relation between two occurrences that act as a single cause, I believe that we have indeed removed the perplexities that Kim mentions, and we even agree with Kim to an extent. We agree, for example, that the occurrences in question are not discernible (or independent) *as causes*. We have not committed ourselves to it, but it is obviously consistent with the causal indiscernibility of a and b that one of these two is "reducible" to or "supervenient" on the other. Thus, when Kim proposes these options for the two occurrences *as against* their "nomic equivalence" or "independence", it does not militate against the claim that they are causally indiscernible. ([Kim89]: 90) As we have been at pains to point out, however, we are agnostic as to whether causal indiscernibility entails complete indiscernibility. For this sense of "independence", then, we do not endorse Kim's claim.

As regards Kim's claim that the two occurrences may not both be sufficient, Kim asks two questions. First:

> Why does the supposed nomological relationship between [a and b] void the claim that this is a case of causal overdetermination? ([Kim89]: 86)

Here, our remarks in **Overdtermination, Single Causes** suffice. Kim makes no attempt to answer the question or to show that the nomic relationship between a and b does not or cannot suffice to rebuff the charge of overdetermination. In the absence of such reasons, then, I take it that our remarks are unchallenged, as are Bennett's.

Kim continues:

[15] Kim himself recognizes that the meanings of "complete" and "independent" are crucial to the discussion, but he nonetheless slips out of making them precise: "The meanings of 'complete' and 'independent' are obviously crucial. I shall not be offering definitions of these terms; rather I shall focus on some specific cases falling under the intended distinctions, with the hope that, in the course of my discussion, reasonably determinate core meanings will emerge that will give the exclusion principle clear and substantial enough content." ([Kim89]: 89)

> Notice the trade-off here: the closer this is to a standard case of overdetermination, the less dependent are the two explanations in relation to each other, and, correlatively, the more one stresses the point that this is not a case of standard overdetermination because of the nomic equivalence between the explanations, the less plausible is one's claim that we have here two distinct and independent explanations. ([Kim89]: 86)

We are in conditional agreement with Kim that the second horn of the implied dilemma applies to *a* and *b*: there is a sense in which they are not distinct or independent. They are neither distinct nor independent as causes—they act as a single cause. It remains unresolved, however, whether they are *ontologically* discernible; if ontological discernibility is equivalent to the sense of "distinct" or "independent" Kim had in mind, then we conditionally reject the second horn of the implied dilemma. The condition is the truth of CE: if it is true, we accept the horn; if it is false, we do not.

5 Conclusions

We have elucidated causal indiscernibility for two contemporary views of causation and shown that the truth-value of CE remains unclear in this light. Still, we may discern two conditions on its truth: that (i) nomic impossibilities do not bear on the identity of a causally embedded occurrence, and that (ii) necessary spatiotemporal coincidence is sufficient for indiscernibility. Conditions (i) and (ii) are necessary and sufficient to justify CE for Davidsonian causal indiscernibility. (i) and (ii) are necessary but insufficient to justify CE for Lewisian causal indiscernibility—there remains the issue of spatiotemporal coincidence within each world. Any defense of CE, then, must justify (i) and (ii). An argument against CE, on the other hand, might undermine (i) or (ii), or it might pursue some other strategy.

In addition, we have situated "acting as a single cause" between joint causation and overdetermination. Where c is an occurrence satisfying the true criteria for being a cause of e, and where 'a' and 'b' refer to occurrences, a and b act as a single cause of e if both a and b are both necessary and sufficient for c.

References

[Ben88] Jonathan Bennett. *Events and their Names.* Hackett Publishers, 1988.

[Ben03] Karen Bennett. Why the exclusion problem seems intractable, and how, just maybe, to tract it. *Nous*, 37(3):471–497, July 2003.

[Ben08] Karen Bennett. Exclusion again. In Jesper Kallestrup Jakob Hohwy, editor, *Beind Reduced: New Essays on Reduction, Explanation, and Causation*, chapter 14, pages 280–306. Oxford University Press, 2008.

[Dav67] Donald Davidson. Causal relations. *The Journal of Philosophy*, 64(21):691–703, November 1967.

[Dav80a] Donald Davidson. Comments and replies. In *Essays on Actions and Events*. Oxford University Press, 1980.

[Dav80b] Donald Davidson. The individuation of events. In *Essays on Actions and Events*, pages 163–180. Clarendon Press, 1980.

[Dav80c] Donald Davidson. Mental events. In *Essays on Actions and Events*, pages 207–227. Clarendon Press, 1980.

[Dre89] Fred Dretske. Reasons and causes. *Philosophical Perspectives*, 3:1–15, 1989.

[Gib75] A. Gibbard. Contingent identity. *Journal of Philosophical Logic*, 4:187–221, 1975.

[Gol69] Alvin I. Goldman. The compatibility of mechanism and purpose. *The Philosophical Review*, 78(4):468–482, October 1969.

[Hit93] Christopher Hitchcock. A generalized probabilistic theory of causal relevance. *Synthese*, 97:335–364, 1993.

[Hit95] Christopher Hitchcock. The mishap at reichenbach fall: Singular vs. general causation. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 78:257–291, 1995.

[Hit96] Christopher Hitchcock. The role of contrast in causal and explanatory claims. *Synthese*, 107:395–419, 1996.

[Hum99] David Hume. *An Enquiry concerning Human Understanding*. Oxford Philosophical Texts. Oxford University Press, 1999.

[Kim89] Jaegwon Kim. Mechanism, purpose, and explanatory exclusion. *Philosophical Perspectives*, 3:77–108, 1989.

[Kim93a] Jaegwon Kim. Events as property exemplifications. In *Supervenience and Mind*, chapter 3, pages 33–53. Cambridge University Press, 1993.

[Kim93b] Jaegwon Kim. Mental causation in a physical world. *Science and Knowledge*, 3:157–176, 1993.

[Kim95] Jaegwon Kim. Mental causation: What? me worry? *Philosophical Issues*, 6:123–151, 1995.

[Kim97] Jaegwon Kim. Does the problem of mental causation generalize? *Proceedings of the Aristotelian Society*, 97:281–297, 1997.

[Kim98] Jaegwon Kim. *Mind in a Physical World.* Representation and Mind. The MIT Press, 1998.

[Lew73] David Lewis. Causation. *The Journal of Philosophy*, 70(17):556–567, October 1973.

[LL87] Ernest Lepore and Barry Loewer. Mind matters. *The Journal of Philosophy*, 84(11):630–642, November 1987.

[Low00] E.J. Lowe. Causal closure principles and emergentism. *Philosophy*, 75(294):571–585, October 2000.

[Mal68] Norman Malcolm. The conceivability of mechanism. *Philosophical Review*, LXXVII, 1968.

[McL85] Brian McLaughlin. Anomalous monism and the irreducibility of the mental. In Ernest Lepore and Brian McLaughlin, editors, *Actions and Events: Perspectives on the Philosophy of Donald Davidson.* Blackwell, 1985.

[McL89] Brian McLaughlin. Type epiphenomenalism, type dualism, and the causal priority of the physical. *Philosophical Perspectives*, 3:109–135, 1989.

[Mel95] D.H. Mellor. *The Facts of Causation.* Routledge Press, 1995.

[Pap93] David Papineau. *Philosophical Naturalism.* Blackwell, 1993.

[Pea79] Christopher Peacocke. *Holistic Explanation: Action, Space, Interpretation.* Clarendon Press, 1979.

[Rea97] M. Rea, editor. *Material Constitution.* Rowman and Littlefield, 1997.

[Rit05] Jack Ritchie. Causal compatibilism: What chance? *Erkenntnis*, 63(1):119–132, July 2005.

[Sos95] Ernest Sosa. Davidson's thinking causes. In John Heil and Alfred Mele, editors, *Mental Causation*, chapter 4, pages 41–50. Clarendon Press, 1995.

[Stu05] Karsten R. Stueber. Mental causation and the paradoxes of explanation. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 122(3):243–277, February 2005.

[VF80] Bas Van Fraassen. *The Scientific Image.* Oxford University Press, 1980.