Response to Ehring's 'Papineau on Causal Asymmetry'
Author(s): D. Papineau
Source: *The British Journal for the Philosophy of Science*, Vol. 39, No. 4 (Dec., 1988), pp. 521-525
Published by: The University of Chicago Press on behalf of The British Society for the Philosophy of Science
Stable URL: https://www.jstor.org/stable/687436
Accessed: 23-07-2026 08:32 UTC

REFERENCES
Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/687436?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

The British Society for the Philosophy of Science, The University of Chicago Press are collaborating with JSTOR to digitize, preserve and extend access to *The British Journal for the Philosophy of Science*

This content downloaded from 194.80.232.28 on Thu, 23 Jul 2026 08:32:26 UTC
All use subject to https://about.jstor.org/terms

Brit. J. Phil. Sci. **39** (1988), 521–525 *Printed in Great Britain*

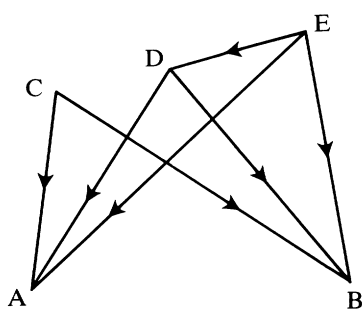
DISCUSSION

Response to Ehring's 'Papineau on Causal Asymmetry'

In 'Papineau on Causal Asymmetry' [1987] Douglas Ehring makes some telling objections to the theory elaborated in my 'Causal Asymmetry' [1985]. It seems to me, however, that his objections count rather more against the letter of my theory than its spirit.

My [1985] paper took off from the fact that 'screening off' is a pervasive feature of the natural world: probabilistic associations between joint effects disappear when we 'control for' (that is, conditionalize on the presence and absence of) their common causes. But in fact this was rather a loose way of describing what is really a more complicated phenomenon. The complication is that joint effects will often have a *number* of common ancestors in their causal history. And then their probabilistic association will only disappear if we control for *all* those common causal ancestors.

The kind of diagrams drawn in multiple regression and path analysis models of causal systems make the point clear.



To get the correlation between A and B to disappear, it's not enough to control for C. You also need to control for D and E. In general, you need to make sure that none of the lines coming into A are left with an uncontrolled-for common source with any of those going into B.

A number of writers, from Reichenbach on, have toyed with the idea that the screening-off asymmetry is constitutive of causal direction, that is, that the

This content downloaded from 194.80.232.28 on Thu, 23 Jul 2026 08:32:26 UTC
All use subject to https://about.jstor.org/terms

difference between causes and effects just *is* the fact that common causes screen off associations between joint effects but not vice versa. In 'Causal Asymmetry' I tried to go a step beyond this by explaining the screening-off asymmetry in terms of something else, and then making that something else constitutive of causal direction. The something else in question was the thesis that the various background conditions in conjunction with which causes fix their various effects are probabilistically independent of each other, while the 'background conditions' in conjunction with with effects 'fix' their various causes are not so independent.[1]

But now, if we consider diagrams like the above, we can see that this too is a rather loose way of putting something that really needs to be more carefully formulated. What we really need is something like the following thesis: in any *well-formed* causal diagram the various background conditions required for the tail of each arrow to fix the head of that arrow are probabilistically independent of each other, while the 'background conditions' required for the head of each arrow to 'fix' the tail of that arrow are not so independent (where what is required for a causal diagram to be 'well-formed' is that, if any two explicitly represented factors or relevant background conditions have a common causal ancestor, then that common causal ancestry should itself be explicitly represented in the diagram). And the overall theory then needs to be that *this* contrast is constitutive of the difference between causes and effects.

It might seem that the mention of common causal ancestry in the above explanation of what makes a diagram well-formed invalidates this theory as a reductive analysis of causal direction. But we can think of things in a way which eliminates this appeal to the concept of causal ancestry. Imagine that we start off with a map containing non-directional lines linking *all* factors which are causally connected to the phenomenon initially under consideration. Our task is to give some of these lines a causal direction. My thesis is that if we can find a way of doing this (in effect, a way of 'cutting out' a causal diagram) such that the background factors at the tail ends of the directional arrows are all robustly[2] probabilistically independent, then this is a well-formed causal diagram, and those arrows will be going from causes to effects, not from effects to causes.[3]

Once we have explained well-formedness in this way, we have a non-circular explication of the notion of causal direction. And then, given this

[1] 'Background conditions' are here to be thought of as factors which aren't explicitly represented in causal diagrams like the above; they are the implicit 'error terms' whose values in particular cases determine whether or not the explicitly considered causes give rise to their effects.

[2] The significance of 'robustly' will be explained below.

[3] It is interesting to note that something like this analysis is implicit in the use of regression and path analysis to build causal models. Regression or path coefficients are only taken to be of causal significance in so far as the exogeneous 'error terms' in the regression equations are probabilistically independent of each other.

This content downloaded from 194.80.232.28 on Thu, 23 Jul 2026 08:32:26 UTC
All use subject to https://about.jstor.org/terms

entitlement to the notion of causal direction, we can go on to recognize that there will exist some diagrams whose arrows all correctly represent directed causal links, but which *don't* have probabilistic independence of background factors at the tail ends of the causal arrows: for a given set of causal arrows may merely give us a *sub*-diagram of a full well-formed diagram, where the full well-formed diagram would show that the sub-diagram had omitted certain lines of common causal ancestry.

Unfortunately, these complications were only gestured at in my original paper. There is only the brief passage on pp. 280–1 (*cf.* Ehring's p. 82). This left the necessary elaboration of my thesis very much implicit. Moreover, this passage also carried the unfortunate suggestion that the only kind of case in which two background conditions will have a common causal ancestry, and so be probabilistically dependent, is where they are joint effects of a common cause—when there are, of course, other possible cases of dependence-inducing common causal ancestry, such as where one background condition is itself a cause of part of the other, or where one background condition is identical with part of the other.

It seems to me that if my overall thesis had been formulated so as to accommodate these complications, then most of Ehring's objections would cease to apply. Ehring's two putative counterexamples on p. 83 involve background conditions that are wholly or partially identical; the same is true in the indeterministic example he discusses at the foot of p. 87; and in the example at the top of p. 84 he postulates pairs of background conditions where part of one condition causes part of the other. All of these are cases where the background conditions have a common causal ancestry (even though they aren't joint effects of a common cause), and so their lack of probabilistic independence wouldn't impugn a more careful formulation of my thesis.

The case Ehring discusses later on p. 84 is slightly more complicated. Here he aims to show that it is perfectly possible to find probabilistic *independence* among the 'inverted background conditions' in conjunction with which common *effects* fix their joint causes, contrary to my claim that such independence is what distinguishes causes from effects. However, he achieves this by supposing that one of the 'background conditions' in question (R) is always either caused by another background condition (A), or by a factor (B) which is negatively correlated with A. In general, such a negative correlation indicates some kind of common prior ancestry; and of course A causing R is itself an instance of common causal ancestry. Once more, I would argue that when these common causal ancestries are taken into account, then the required probabilistic dependence of 'inverted background factors' will reappear.

This example also raises another difficulty, however. Even if the desired probabilistic dependence of 'inverted background conditions' can be restored by taking into account common causal ancestry, it remains the case that, prior

This content downloaded from 194.80.232.28 on Thu, 23 Jul 2026 08:32:26 UTC
All use subject to https://about.jstor.org/terms

to taking that common causal ancestry into account, we have an causal diagram with the background conditions at the tail end of the arrows probabilistically independent, even though the arrows in fact go from a common effect to its joint causes. So why isn't this a well-formed causal diagram, given my earlier remarks, and why doesn't it then refute my claim that the probabilistically independent tails of arrows in well-formed diagrams are always causes? Here I appeal to the requirement of *robust* probabilistic independence to which I drew attention in footnote 2. It is true that A and R appear probabilistically independent if we don't control for any further factors. But in practice, I claim, their 'real' probabilistic dependence would reappear once we did make explicit the relevant common causal ancestry. The background factors in a well-formed causal diagram, by contrast, are required to be robustly independent, in the sense that they have to remain probabilistically independent whatever further factors we being into explicit account.

Actually, whether the requisite probabilistic dependence between A and R would reappear in *Ehring's* example, as opposed to in a real case of similar structure, is a moot point: Ehring places such strict requirements on the correlations between A, R, and B that some of the relevant conditional probabilities become undefined. But I would argue that this makes Ehring's case unrealistic. In practical cases there always remain 'error terms' that have not been brought into explicit account, and the resulting 'unexplained variance' keeps the relevant conditional probabilities defined.

This raises a general point about my theory and Ehring's criticisms of it. Clearly, my theory isn't at all plausible as conceptual analysis: even before we go into details, it's obvious that there's nothing about probabilities or background conditions in our everyday concept of causal direction. However, my aim isn't conceptual analysis, but metaphysical reduction: I'm trying to identify the essential characteristics of the direction we intuitively recognize as the cause-effect direction. So examples designed to show that it's conceptually possible that a cause-effect relationship might lack those characteristics will cut no ice against my theory. I don't claim that such cases are conceptually impossible. My claim is simply that my theory identifies the real nature of the cause-effect relationship.

This last point is relevant to the very first objection Ehring brings to my theory (top of p. 83). Here he postulates a pair of joint effects of a common cause, whose respective background causal factors themselves have a common cause, and so on *ad infinitum*. Given this set-up, it would be impossible ever to take enough factors into explicit account to leave the remaining 'error terms' probabilistically independent, and so impossible to construct the kind of 'well-formed diagram' required for my theory to be applicable. Ehring's thought is that this clearly wouldn't stop the causes in the postulated set-up from *being* causes; and so causal direction must be something different from what my theory says. My response is to concede, for the sake of the argument,

This content downloaded from 194.80.232.28 on Thu, 23 Jul 2026 08:32:26 UTC
All use subject to https://about.jstor.org/terms

that such infinitely receding structures of common causes might be conceivable, but to deny that such structures can ever really occur.

My theory is based on the assumption that the screening-off asymmetry is the key to the nature of causal direction. The screening-off asymmetry is that associations between joint effects are screened off by their common causal ancestry, but associations between joint causes aren't screened off by their common descendants. Ehring's postulated case has joint effects whose association can't be screened off by controlling for their common ancestry. Maybe there is a sense in which we can *imagine* such a case. But if I am right about the significance of the screening-off asymmetry, then the case we are imagining is one which lacks the structure on which causal asymmetry in fact rests, and so isn't really possible.[4]

D. PAPINEAU
Cambridge, U.K.

[4] I haven't discussed Ehring's final argument (p. 86), which is a consequence of some sloppy terminology on p. 284 of my original article. Ehring notes an ambiguity in my saying that some factor is required 'to ensure the cancer did come from the smoking and not from the asbestos'; and he points out that the surrounding argument doesn't make much sense on either reading. But I'm afraid I had in mind a yet third reading: I meant that it was only when conjoined with the factor in question that the cancer became a sufficient condition for smoking. I still think that the argument is all right on this reading.

REFERENCES

EHRING, DOUGLAS [1987]: 'Papineau on Causal Asymmetry', *British Journal for the Philosophy of Science*, **38**, pp. 81–7.

PAPINEAU, DAVID [1985]: 'Causal Asymmetry', *British Journal for the Philosophy of Science*, **36**, pp. 273–89.

This content downloaded from 194.80.232.28 on Thu, 23 Jul 2026 08:32:26 UTC
All use subject to https://about.jstor.org/terms