Nancy Cartwright

Comments on Longworth and Weber

Article (Accepted version)

Original citation:
Cartwright, Nancy (2010) Comments on Longworth and Weber. Analysis, 70 (2). pp. 325-330.
ISSN 0003-2638
DOI: 10.1093/analys/anp158

Comments on Longworth and Weber discussion

As Francis Longworth discusses, *HC&UT* argues that there is no such thing as *the* causal relation nor a handful of causal relations nor even a truck load. Rather, there is only a seemingly endless array of relations, called 'thick' relations in the book, that may be loosely grouped under the label 'causal', with a vast variety of different clusterings for different purposes. So I do not see Hitchcock's question as posing an objection. Once we have served our purpose by labelling a relation 'causal', there may well be nothing added, helpful to this purpose, by noting further that the relation is one of feeding gasoline to the carburettor rather than, say, stuffing it in or making it available. For instance, if our purpose is to predict how probabilistic changes propagate across a set of quantities supposing the probability of one is changed 'surgically', this purpose is served once we can point to the probability-change-making relations among these quantities and label them as 'causal' in the Bayes-nets sense.

My purpose in labelling a very great many relations as 'causal' is to counter current-day 'Humeans', like David Lewis, John Earman and Bas van Fraassen. These philosophers are enamoured of some one or another special set of features in nature, features that are conceived of as 'inert', 'non-active', 'temporally and spatially local' or whatever. A great many features I see around me – and that appear to be essential in every successful scientific intervention I have studied – are missing from their special sets. These philosophers – the 'Humeans' – are the ones who make the causal/non-causal distinction, with their favoured concepts labelled 'non-causal'. In maintaining that pushing, pinching, compressing, repelling, attracting and so forth are causal, I imply that these are not among the favoured sets of the Humeans and they are not reducible to features in these sets. And of course I also maintain that they are every bit as much in the world, and play as great a role in real science, as the 'non-causal' features the Humeans admire.

So, thick descriptions of relations labelled 'causal' need not be causal-content increasing. Indeed they can be causal-content decreasing. Consider 'the pesticide smothers eel worms', which is a possible thick description for an arrow in a causal-Bayes-net of Judea Pearl representing 'the pesticide causes eelworm death'. The arrow is appropriate only if smothering is part of a set of relations that together satisfy the Bayes-nets axioms. But being part of such a set is not implied by 'smothers'. This fact may make trouble for Longworth's own proposal, assuming, as I would, that being correctly representable by a causal-Bayes-net arrow should be among Longworth's right-hand-side disjuncts.

I hope that Longworth can steer around this kind of problem, however, because the scheme he offers can make causality a far more useful concept than I fear it to be. Consider: If thick descriptions are not composites of 'causes' plus some non-causal description there may well be criteria for the application of the thick label that don't include the characterizing features pointed to in labelling the relation 'causal'. If in turn the thick description implies these features, we then have an independent way to identify when these features obtain. This is a powerful tool. Consider for instance how useful it would be to have independently accessible criteria for when a set of relations satisfies the Bayes-nets axioms or transmits a mark or reliably produces a targeted outcome.

The possibility of inferring from thick descriptions to 'causes' is independent of the other half of Longworth's scheme, which takes causality to be analysable as a (possibly very long) disjunction. Here I have a weaker view: Family resemblance is the best we can hope for.[1] Nor do I see why we need the stronger claim. If the upward implications from thick descriptions to features characterizing causal laws work, Longworth's answer to Hitchcock goes through whether or not a proper 'if-and-only-if' can be provided. And, as I just argued,

[1] Hence I don't really hold with Longworth's interpretation (iii) of me.

the upwards implications would be a boon whether or not Hitchcock's question raises problems for my claims about thick relations.

With respect to 'if and only if', I think there's no point in doing it unless it can be done properly. There I'm sceptical. Hall's two concepts are not two concepts but two loose groupings, as are the labels in Longworth's (DTC_1). Consider just the probabilistic theory of causality. *HC&UT* argues that the correct expression of it depends on the exact form of the causal-probabilistic laws at work as well as a variety of facts about the population. Nor should this be surprising. Causation is a concept from ordinary life – what Otto Neurath called a 'ballungen' concept:[2] a vague, rough-edged congestion of ideas and criteria that serves a lot of different purposes at once and often none of them very precisely. As Stuart Hampshire maintained, such concepts should not be subject to more rigour than they can bear.

Yet to figure in science they must be rigorous. Even narrowing consideration to scientific contexts, though, I would not expect 'if-and-only-if' to be possible. As Sophia Efstathiou stresses,[3] when ordinary concepts are brought into science they get transformed and transformed into concepts that do different but very specific jobs in different sciences in different ways. Quantum field theory requires a different concept of causation than bicycle makers do, which is different again from concepts needed in various social sciences. Even within economics a number of different concepts are available, and useful, as *HC&UT* illustrates. There's no reason to think that this array of scientifically-transformed causal concepts is finite or fixed or has any system to it.

As to Weber's comments, I agree with all, except his interpretation of me and his enthusiasm for IARC. I am less enthusiastic not because I advocate less diversity, but more. My distinction between methods that clinch conclusions and those that merely vouch for them is part of a larger study of evidence-ranking and conclusion-grading schemes.[4] There is now an array of these, all looking much like IARC's, though often less nuanced. I argue that we need accounts of 1)the quality of evidence – how likely is the evidence claim to be true, 2)relevance – in what ways and how strongly does it bear on the conclusion, and 3)how to evaluate conclusions in light of all evidence. Most guides focus on 1), ignore 2)[5] and give flawed advice about 3).

In IARC we see 1) in the assessment of individual studies ("Each [epidemiological] study is assessed according to three criteria…."; for experimental animals, "Conclusion (1) is drawn if there are *high quality* studies…").[6] 3) is done in two steps: First assign *sufficient*, *limited*, *inadequate* or *lack* labels for each of the three categories (epidemiological, animal studies, mechanistic); second, draw the conclusion labelled 'group x' according to the rules Weber sketches. So we can conclude "carcinogenic to humans" if there is sufficient epidemiological evidence for carcinogenicity in humans and we can conclude this if "a positive relationship has been observed in studies in which chance, bias and confounding could be ruled out with reasonable confidence." Thus unbiased, etc. epidemiological studies trump the rest: Positive results from these yield a positive conclusion even if evidence from the other modes is inadequate or negative.

---

[2] For further discussion of this see Cartwright, Cat, Fleck and Uebel (l996).
[3] See Efstathiou (2009) on what she calls 'founded' concepts.
[4] See my website (http://personal.lse.ac.uk/cartwrig/Default.htm) for my own papers on these topics and references therein to others.
[5] For instance, relevance, as opposed to quality and methods for drawing overall conclusions, is not explicitly addressed in the IARC scheme. But looking at the scheme, as Weber describes it, it appears that the relevance of positive epidemiological results on *some* human population to 'is carcinogenic to humans' is taken for granted (I think fallaciously), whereas the relevance of animal results requires a hypothesis about shared mechanisms.
[6] IARC (2006).

Why? The only sensible reason I have come up with is that of chapter I.3: These unbiased (etc.) studies are **clinchers** – *given the background assumptions*,[7] a probabilistic difference in the occurrence of cancer between exposed and unexposed groups deductively implies that the agent caused cancer in the study population.[8] The importance of discussing clinching is not, as Weber suggests, to argue that a single study is sufficient but the reverse. Why prefer these studies? 'Because they are unbiased (etc.)'. What's so good about that? In the end, I take it, the answer is that lack of bias (etc.) is enough to guarantee the conclusion. But now it is hard to see why these studies trump others. Deductive conclusions are only as sure as their premises. Premises are never certain.[9] So the conclusions aren't either. We give a good shot at ensuring the premises. But why have more confidence that we have succeeded than we have in arguments that may have equally – or even more – secure premises but not such a tight fit between premises and conclusion? From this perspective the most sensible procedure seems to be to look at all the studies together and give an overall evaluation, as best possible, not to suppose that following a recipe to reduce bias (etc) will always produce more reliable results than studies that rely on mechanisms or animal models or even statistical studies with some bias.

Weber also takes my connection between types of systems or types of conclusions on the one hand and types of evidence on the other to suggest conclusions be drawn from one type of study alone. In a sense I maintain the reverse. All methods require background assumptions; when the assumptions fail, the methods cannot be taken to be reliable. Yet we often cannot tell whether the assumptions are met or not. It is good practice then to multiply studies in order to control for error in assumptions, and in the end to make an overall judgment based on the pattern of results.

Beware though. The reliability of method depends on the type of conclusion drawn. So Weber's concern that differentiating kinds of conclusions, as I insist we must, restricts the methods we can use to test them, is a correct one. Indeed it is one of the central lessons I urge, and one of great practical significance. We need to be clear about *exactly what kind of causal conclusion* we are entitled to draw from our studies; and not out of philosophical purity but because the inferences we can go on to draw as we imbed that conclusion in other contexts, combining it with other knowledge, depend intimately on what causal claim we have established in the first place. Indeed, this is one of my worries about current causal investigations in the biomedical and social sciences: They are often sloppy about what the causal conclusion drawn claims.[10]

What then of the scatter of methods Weber describes from IARC – epidemiological studies, studies on model animals, studies of mechanisms? These can properly be used in a complex inference to establish not causal laws (which in my account are always population-specific) but more abstract capacity claims: claims about a stable contribution the cause can make. It is a complicated story about what is involved in a capacity claim and how such claims can be justified.[11] But epidemiological methods and studies of mechanisms can play an important role.

---

[7] Including common assumptions about the relation between causes and probabilities as well as the context-local assumptions that the two groups have the same distribution of other causally relevant features (unbiased) and that there is a genuine probabilistic difference in outcome (the difference in frequency is not a result of chance).

[8] That's because, given the assumptions, the only possible explanation for the difference is that the agent caused cancers in some members of the exposed population.

[9] Including the broad metaphysical premise required in these arguments, which I discuss in *HC&UT*, that every probabilistic difference requires a causal explanation.

[10] Sometimes this is 'remedied' by reporting precise statistical results. This can be an even worse practice since these results are often only useful when interpreted causally – and that is the interpretation they are given. But in this case the causal interpretation, being less explicit, is even less rigorously reviewed.

[11] For further discussion see Cartwright (2007) and (1989) and articles on my website on causal powers.

Finally, Weber urges attention to simulations and experiments in economics. We are again here in agreement: This is an important topic. I discuss only thought experiments, in a chapter on 'idealized' models, though many of the cautions about 'overconstraining' restrictions in the set-ups apply to actual experiments as well. Other than that, I leave experiments and simulations to other colleagues, including Weber himself, who have more interesting things to say than I would.[12]

References

Alexandrova, Anna (2006), 'Connecting Economic Models to the Real World', *Philosophy of the Social Sciences*, 36: 173-192.

---- (2007) 'Making Models Count', *Philosophy of Science* 75: 383-404.

Cartwright, Nancy (1989). *Nature's Capacities and their Measurement,* New York: Oxford University Press.

---- (2007). *Causal Powers: What Are They? Why Do We Need Them? What Can and Cannot be Done with Them?* Contingency and Dissent in Science Series, London: Centre for Philosophy of Natural and Social Science, LSE.

Cartwright Nancy, Jordi Cat, Lola Fleck, Thomas E Uebel (1996). *Otto Neurath: Philosophy between Science and Politics*, New York: Cambridge University Press, Ideas in Context Series No. 38.

Efstathiou, Sophia (2009). *The Use of 'Race' as a Variable in Biomedical Research*, PhD Dissertation, University of California San Diego, UMI Press, 2009.

Grüne-Yanoff, Till (2006). '*Learning from Economic Models*', Paper presented at the 20th Biennial Meeting of the Philosophy of Science Association, Vancouver.

Guala, Francesco (2005). *The Methodology of Experimental Economics*, Cambridge: Cambridge University Press.

Hausman, Daniel M. (1992). *The Inexact but Separate Science of Economics*, Cambridge: Cambridge University Press.

Hendry, David (2001). *Causality in Macroeconomics*, Cambridge University Press.
--- (2004). 'Causality and Exogeneity in Non-stationary Economic Time-Series', *Causality: Metaphysics and Methods* Technical Report CTR 18-04, CPNSS, London School of Economics.

[12] Cf. Guala (2005), Alexandrova (2006), (2007); Gruene-Yanoff (2006); Humphreys (2009); Reiss (2007), Morgan (2001).

Hoover, Kevin (2001). *Causality in Macroeconomics*, Cambridge University Press, Cambridge.

Humphreys, Paul (2009). 'Computational Economics', *Oxford Handbook on Philosophy and Economics*, Harold Kincaid and Don Ross (eds). Oxford: Oxford University Press.

IARC (2006). Preamble to the IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Available at http://monographs.iarc.fr Last accessed, September 4, 2009.

Morgan, Mary (2001). 'Models, Stories and the Economic World', *Journal of Economic Methodology 8;3*: 361-384.

Reiss, Julian (2007). *Error in Economics: The Methodology of Evidence-Based Economics*, London: Routledge.

Woodward, James (2003). *Making Things Happen*, Oxford University Press.