DOI: 10.1017/psa.2022.99
This is a manuscript accepted for publication in *Philosophy of Science*.
This version may be subject to change during the production process.

Review of James Woodward's `Causation with a Human Face`

H.K. Andersen

Department of Philosophy, Simon Fraser University 8888 University Drive, Burnaby, BC V5A 1S6, Canada Email: handerse@sfu.ca

It would be hard to overestimate the amount of progress on causation and causal explanation since Woodward's first book, `Making Things Happen` (2005). It is unusual in philosophy to pronounce so confidently that a discussion has not merely changed, but genuinely progressed. It's easy to say that this book has been long anticipated, will be widely read, and will shape the discussion for years to come. Given that, this is a perhaps surprising direction toward which to turn current discussions of causation, which have been moving more in the direction of formal methods, machine learning, and automated causal discovery. This book goes almost completely the other direction.

`Causation with a Human Face` is a book-length treatment of topics Woodward has been urging philosophers of science toward for some time. It centers the ideas that there is invaluable work to be done in understanding what causation is, by looking at why humans are interested in it, how human causal cognition develops, and what we can learn about causation by considering psychological or cognitive research on it. In "Agency and Interventionist Theories" (2009), for example, Woodward is already pressing us to consider the close connection between humans as agents and a non-anthropocentric account of causation. This book develops those ideas, systematizing a variety of distinctions and illustrating how to engage in this new kind of project by doing so in great detail.

The title is aptly chosen for the emphases that Woodward places on the human-centered and -limited aspects of causation. Rather than focusing on, for example, what can be done in machine learning to reduce or eliminate human peccaries from causal modelling, Woodward turns the other direction, to more fully and completely embrace what is distinctively human about causal

knowledge. He uses terms like causal reasoning, cognition, and judgement, that emphasize dual normative and descriptive aspects. Woodward offers the book as developing what he takes to be a new kind of project that stands in contrast to what he conceives of as traditional epistemological and metaphysical projects. His approach to human causal reasoning situates it as the source of norms on which we rely in more formalized causal reasoning, such as contemporary causal modeling. Criteria such as proportionality and invariance are a refined ancestor of the norms on which humans already rely implicitly.

There are certain criticisms that Woodward has fielded repeatedly from critics who find nonreductive accounts circular, and want a more metaphysical view of what causation is, rather than how we identify it. Such critics will not find answers here. Woodward has repeatedly insisted that is not his project; it is a project, but not the one in which he is engaged. This book vindicates his insistence on his project, as the potential pay-off for refining formal inferences by drawing on richer empirical work on causal cognition becomes clear.

With all the psychological research he covers, it would be easy to think this is some kind of naturalizing project, reducing justification in causal modeling to psychological circumstances, or to think it is naturalistic, focused only on how we actually happen to reason, without implications for how we ought to reason. His project is neither of these two things. Instead, in my view, it is structurally analogous to work by Burge, such as "Perceptual Entitlement" (2003). Woodward situates normative inferential considerations for causal inference as emerging from widespread patterns of informal causal reasoning: the most refined definitions of proportionality, for example, can be found in rough shape in empirical psychology's study of human causal reasoning. The goal is not to eliminate or reduce away the rational norms involved in more formalized causal modeling. It is instead to situate and explains the norms we do use, such as modularity or proportionality, as progressive refinements of implicit rules already discernible in causal cognition. The normative and evaluative features of causal reasoning remain, grown out of and shaped by causal cognition.

The first chapter lays out his synoptic view of how normative claims about causal modelling fit into descriptive projects like empirical psychology of human causal cognition. The norms in question are theoretical, methodological, or epistemological in character: for example,

https://doi.org/10.1017/psa.2022.99 Published online by Cambridge University Press

modularity or invariance are such norms (the norms are not human in the sense of ethical or values-oriented). Descriptive projects includes work by psychologists such as Gopnik, Cheng, Lien, and collaborators, on human causal cognition under a variety of circumstances and at a variety of levels of cognition. Woodward's aim in this chapter is to convince the reader that these two approaches, normative and descriptive, are closely connected, and that there are more connections from the descriptive to the normative than have yet been fully explored.

Woodward offers what he calls a functional account of causation. In order to understand causation, we should consider what function(s) causal reasoning plays in human thinking and acting, a question addressed in empirical psychological work. “If we are to think about causal cognition in functional terms, then (I will argue) we need accurate empirical information about how adult humans (and I will suggest, other subjects such as small children) judge, learn, and reason causally--what they do. We also need information about why they judge, etc., as they do--what factors influence judgment and reasoning--and about how these make for success or failure in their reasoning processes.” (19).

The normative considerations here are not merely inference rules, but more like "the goal or purposes served by some features of causal thinking" (19). Functional accounts look at what we use causation for, in situated, concrete ways. This is similar to pragmatism about causation (e.g., Woodward 2023).

Woodward makes the case that statistics and machine learning also need a foundation in or connection to human causal judgments, because that is how we can say that such work is about causation in the first place. What makes the patterns detected by a sophisticated machine learning program `causal` structure, rather than some other kind of structure? Woodward offers a characteristic deflationary and non-reductive view about causation here. We don't have to give a set of necessary and sufficient conditions for causation, but there must be some way to differentiate why only some statistical relations highlight causal structure.

Chapter 3 gets into three methodological approaches that could be used to investigate causal cognition. The chapter is structured in subsections that don't map easily onto existing discussions in metaphysics of causation. His concern with 'armchair philosophy' involves intuitions as they apply to cases, such as Gettier cases or double pre-emption. He claims intuitions are either a

https://doi.org/10.1017/psa.2022.99 Published online by Cambridge University Press

judgment we make ourselves about what the 'right' causal structure is, or about what judgments others would make. He notes that these intuitions and judgements have no calibration method, lacking standards for accuracy or reliability.

In section 3.3, he gives two possibilities about how they might be considered relevant to a theory of causation: intuition and judgement can be relevant for "an account of what causation 'is' and an account of 'our concept' of causation" (140). This section draws what strikes me as an unhelpful dichotomy: even if one is looking for an account of what causation is, then this cannot be an independent project from what our concept of causation is. No account of what causation 'is' could differ so much from our concept of causation that these are capable of conflicting to such a high degree. The framing of the contrast is drawn from X phi, and ill-fits the causation literature. Separating accounts of what causation 'is' (with scare quotes) from accounts of our concept of causation is a straw man target. He does conclude this chapter with some helpful remarks on how to approach empirical research on causal cognition, where the goal of such an account is explaining the success of our causal reasoning. It is indeed a striking feature of causal cognition that, even though we are sometimes wrong, we get it right as often as do.

Chapter 4 lays out some empirical work on causal cognition and connects it to descriptive and normative work on causation. It illustrates how such work should be done, not by describing it from a high vantage point, as in chapter 1, but by showing how it proceeds by doing it. The chapter considers what it takes to have causal representations: not merely knowledge of relations that are causal, but representations of those as causal. He provides six criteria for causal representation that are, in my view, some of the most fruitful parts of the book. He includes the role of agency in developing causal cognition, as well as the role of causal cognition in agency, thus standing apart from many contemporary discussions of causation in the empirically oriented literature by taking agency seriously.

In chapters 5 through 8 Woodward demonstrates how to use the interconnections between normative causal reasoning and descriptive empirical work to illuminate distinctions within causation. These chapters lay out key ideas in normative causal reasoning, especially proportionality and invariance, and then shows how these fit into the landscape of empirical work on casual cognition. He calls these distinctions within causation. Rather than being about

https://doi.org/10.1017/psa.2022.99 Published online by Cambridge University Press

distinguishing causation from anything that isn't causation, or distinguishing when a causal relation exists as a binary yes/no, these distinctions within causation allow for a more fine-grained way of speaking about causal relations.

Woodward frames the investigation into invariance as partially based on the frequent success of our causal reasoning "...because the search for invariant relationships of various sorts makes normative sense, it is often reasonable to expect that, as an empirical matter, the identification of such relationships and reasoning in terms of them will play an important role in people's causal cognition." (227).

He goes through (at least) 10 different ways in which a causal relationship could be invariant to different degrees. There isn't space to do justice to this here, but the combination of scope and detailed is impressive. This chapter concludes by noting how invariance involves our limitations as humans, both epistemic and calculational. Tracking invariance helps improve success in the face of these limitations, and invariance itself is shaped by our needs in navigating the world with these limitations, using causal relationships.

I'll add two critical points of sorts. The first is not a shortcoming of the book so much as a lacuna worth pointing to. It is about the potential for human reasoning to be done poorly, in systematic ways; this is not limited to causal reasoning, but neither is causal reasoning exempt. Woodward's emphasis on the role of empirical work on causal reasoning for refining normative measures such as invariance is extremely interesting, especially insofar as it provides an independent avenue for development of formal techniques compared to e.g. machine learning. No one, including Woodward, wants to endorse the naturalistic fallacy that because humans `do` reason certain ways, we `ought to` reason in those ways. As he notes, our success in causal cognition is a solid assumption in many circumstances. Yet examination of the inferences embedded in causal representations can also be furthered by considering factors involved in producing systematic *failures* of causal reasoning. Some things in human reasoning go wrong very consistently, can be consistently evoked, and may even be named fallacies (such as Post Hoc, Ergo Propter Hoc). While it may go beyond the bounds of this book, Woodward's project would be complemented by an examination of factors that systematically mislead or trick causal cognition.

https://doi.org/10.1017/psa.2022.99 Published online by Cambridge University Press

The second, as noted above, is that there is some potential straw manning with respect to the critical portions of the book directed at metaphysics of causation. It is difficult to reconstruct what views Woodward is criticizing in some places, and he does not often identify examples with citations. Even as I also disagree with some work in contemporary metaphysics of causation, it was hard to see the shape of that discussion in the criticisms levied against it. Some approaches relying in extremely different ways on a priori or non-empirical methods, and invoking diametrically opposed intuitions about e.g. laws and necessity, were lumped together.

On a final note, this book is quite readable; Woodward has taken a more relaxed and engaging style than in earlier work. This conversational tone makes the sometimes dense sections more absorbing, and helps the highly detailed deep dives into empirical literature hang together with the sweepingly synoptic overviews as a single view that is systematically complete.

References:

Burge, Tyler. 2003. "Perceptual entitlement." Philosophy and Phenomenological Research, 67(3): 503-548.

Woodward, James. 2005. Making things happen: A theory of causal explanation. Oxford: Oxford University Press.

Woodward, James. 2009. "Agency and interventionist theories." The Oxford Handbook of Causation, eds. Helen Beebee, Christopher Hitchcock, and Peter Menzies. Oxford: Oxford University Press.

Woodward, James. 2023. "Sketch of some themes for a pragmatist philosophy of science." The Pragmatist Challenge: pragmatist metaphysics for philosophy of science, eds. H.K. Andersen and Sandra D. Mitchell. Oxford: Oxford University Press.

https://doi.org/10.1017/psa.2022.99 Published online by Cambridge University Press