



Causation, from a human point of view

James Woodward: Causation with a human face: normative theory and descriptive psychology. Oxford: Oxford University Press, 432 pp, £64 HB

Violetta Manola¹ · Stathis Psillos¹

Accepted: 10 July 2023 / Published online: 2 August 2023
© The Author(s) 2023

In his fascinating new book, *Causation with a Human Face* (henceforth *CHF*), James Woodward develops a novel idea about how to think about and study causation by starting from how we humans reason causally and then moving to accounting for the normative aspects of causation, thereby offering a human-centered shift in perspective on these issues. What renders this book a must-read, and an exceptional one of its kind, is primarily the enlightening and multi-faceted way in which Woodward proceeds in actually implementing these novel ideas; in short, his method. He is not just describing a new approach from a panoptic philosophical–theoretical perspective but acts as an “epistemic or conceptual engineer” (30), who views causal thinking as a “tool or technology for achieving certain goals” aiming to systematically connect different aspects of causation with their functional roles in human reasoning. *CHF* acts more as a methodological guide for this new “functional” perspective, rather than as a typical theoretical monograph.

This perspective is even more impressive, given that the main ambition of the book is to overcome the dichotomy between the descriptive and the normative, by connecting the descriptive aspect of causal reasoning, i.e., how humans actually make causal judgements, with its normative aspects, i.e., “how we ought to reason causally”. Hence, it connects facts derived from empirical research on causal cognition with a normative evaluation of the right kind of reasoning, measuring the success or failure of our causal cognition with respect to its function. Woodward’s framework is thus neither an alternative account of what causation *is*, nor a normative assessment of our causal inference methods; it is also not simply an epistemological analysis of causal cognition. *CHF* masterfully incorporates all

✉ Violetta Manola
manola.violetta@gmail.com

Stathis Psillos
psillos@phs.uoa.gr

¹ Department of History and Philosophy of Science, National and Kapodistrian University of Athens, Athens, Greece

of the work on causation, from empirical research on causal cognition to formal/philosophical accounts of causation, without disposing any of it, situating it though under the ‘umbrella’ of the “functional account of causation”. Results from empirical studies on causal cognition, formal approaches like causal modeling and causal inference models in statistics, philosophical theories of causation are all being examined in detail through the lens of the “functional” account.

Another dichotomy that *CHF*’s functionalism aims to overcome is between epistemology and metaphysics. Without conflating the two different enterprises, Woodward stresses that the epistemology and the metaphysics of causation should constrain each other. As he put it: “*how* we find out about causal relations can tell us something about *what* those relationships involve”. “Conversely”, he adds, “a satisfactory account of what causal relations involve can help us to understand how our procedures for finding out about them succeed to the extent that they do” (10). In any case, Woodward subscribes to the recent tendency toward minimal metaphysical assumptions. What he calls “minimalist metaphysics” dispenses with all philosophical theories of the truth-makers of causal claims, be they neo-Aristotelian or neo-Humean, on the grounds that they go far beyond what science tells us about nature, themselves implicating a sort of *sui generis* “metaphysical explanation” (7).

What exactly is, then, Woodward’s functionalism? This is summarized by the idea that “causal cognition or thinking in causal terms is useful or functional in successfully serving various goals and purposes that we human beings (and perhaps other creatures) have” (28). This human goal-theoretic account of causation, which is tantamount to the claim that understanding causation involves understanding the conceptions *we* humans have developed and the practices of causal judgment *we* share and what functions these have for *us*, might create the impression that causation, as it is in the world, is mind-dependent. Woodward takes pains to ward off this misunderstanding. His minimalist metaphysics is defended as “minimal realism”; *realism* in the sense that it takes it that causal relations exist “out there” in the world; and *minimal* in that it refrains from telling any allegedly deeper story as to what these causal relations *really* are. The whole project of *CHF* is predicated on the claim that successful causal reasoning *requires* that certain relations are present in nature. So, a functional approach to causation is characterized by the mapping of various features of causal cognition onto goals that we have (typically, manipulation, prediction, and control), whereas its normative evaluation is actually an assessment of whether and to what extent those features of causal cognition succeed in achieving their goals. This understanding contrasts with more traditional metaphysically driven philosophical or formal accounts of normative justification that focus on the discovery or grasping of the “true” nature of causation.

In order to see how all this plays out, let us take a quick look at the old dictum: Correlation does not imply causation. The functionalist approach, by examining issues such as why our concept of causation implies this dictum, is interested in the issue of why this dictum is valued. For instance, on the interventionist framework adopted by Woodward, correlation is distinguished from causation in that it is the search for stable relations that is valued, that is of relations “potentially exploitable for purposes of manipulation and control” (30).

Before we go on, let us take a quick look at the individual chapters. In Chapter 1, Woodward presents the general framework of the book, he introduces the reader to his functional account on causation and sketches how the normative and the descriptive are interconnected and can mutually inform one another.

In Chapter 2, he proceeds by evaluating different conceptions and theories of causation, such as the regularity theory, probabilistic theories, counterfactual conceptions or causal models based on graphs and structural equations. He does that by focusing on whether and to what extent they capture how humans actually make causal judgments and carefully extracts those aspects of the causal theories that match, and correspond to, our ability to reason causally and make causal inferences.

Two important features of this chapter need highlighting. The first is that he explores different accounts of causation, difference-making and geometrical/mechanical ones, by completely disregarding the debate on which theory is more fundamental in grasping the alleged hierarchical causal structure of the world or similarly the debate on whether counterfactuals are more fundamental than causation or the other way around. His selective shift of focus on the practice of our causal reasoning makes way for the integration of different causal theories and concepts under the common functional umbrella. In trying to understand to what extent and in what ways geometrical/mechanical causal theories such as process accounts by Dowe and Salmon capture features of our causal reasoning, he highlights how this is often integrated with difference-making accounts in our causal reasoning practices. He suggests, for example, that humans often use geometrical/mechanical information to infer difference-making relations. Thinking through the functional framework suggests using geometrical/mechanical considerations, not in order “to define or characterize what it is for a relationship to be causal” (104), but instead in order to understand its “relationship with difference-making relations, on the assumption that the former can convey information about the latter” (104).

The second feature of Chapter 2 that needs highlighting is Woodward’s fruitful way of looking at the distinction between token/type causal claims. Again, he dismisses the common debate in philosophy of causation on which causal claims (token or type) are more scientifically or metaphysically fundamental, and instead tries to understand what function each one has in human causal reasoning. In doing so, he finds out that actual/token causal claims follow an opposite inference direction than type causal claims: the first goes backward from effect-to-cause, while the latter follows a forward reasoning from cause-to-effect (see 109). He relates this different inferential direction with a potential different function that token causal claims might have. For example, he suggests that token causal claims might be used when we want to assign responsibility for an action or event: “when we make an actual causal judgment regarding some occurrence, we attempt to trace back the causal history or etiology of the occurrence with an eye to finding some action or natural occurrence that we can hold “responsible” for the occurrence” (112).

Chapter 3 is the least readable and flowing chapter of the book. In it, Woodward explores the role of philosophy, armchair and experimental, and empirical research as methods for studying causal cognition. Although philosophical methods involving intuition and judgment about cases are of limited usefulness in their alleged roles as sources of information about the nature or concept of causation, they can

nonetheless provide *some* information or evidence about some aspects of causal reasoning and cognition. However, in that respect, “there is nothing special about the intuitions or reports of judgments about cases produced in philosophical contexts” (132). Empirical methods are often better as sources of information about causal cognition, and one reason is that they can construct non-verbal cases where subjects actually perform causal tasks, rather than just verbally respond to verbally formulated cases. It seems important for Woodward in this chapter to critically evaluate what philosophical methods can and cannot establish regarding the discussion on causation. Toward the end of the chapter, he discusses ways of explaining the relative success/failure of our causal reasoning by looking at empirical work and is concerned again with the normative assessment of the descriptive, as well as their interconnections. He argues that “empirical results (whether they derive from people’s intuitive judgments understood as a basis for claims about how people judge and think or derive from some other source) are not, in themselves, “evidence” for any particular normative theory”, but at the same time “some normative theories would appear to lack obvious motivation if as an empirical matter people never reasoned in accord with them” (162).

In Chapter 4, Woodward presents a number of empirical results relevant to causal learning and causal representation and tries “to connect these results to various normative and descriptive theories of causation, including associationist, counterfactual, interventionist, and process-oriented accounts” (169). The context is given through Woodward’s list of criteria through which to evaluate the strength of causal representation, which amounts to a subject representing a relationship *as* causal (see 170ff).

Chapter 5 is where Woodward elaborates on his choice of “invariance” as the main normative notion of his account, where invariance is tantamount to “sameness of operation or applicability of a relationship across changes” (262). He shows why invariance as a notion is useful in causal analysis, not only as a minimal condition for identifying a relation *as* causal (yes or no answer), but also as a tool to measure the relative strength of a causal relation (degrees, grades of invariance). He explores various ways in which causal relations differ with respect to the sorts of changes over which they are invariant (for example by changing their background conditions or by intervening on them).

In Chapter 6, Woodward applies the normative notion of invariance to common sense causal thinking, while in Chapter 7, he looks at empirical work on the role of invariance in causal reasoning. Both of these chapters can be seen as a way to evaluate and test this normative notion across different fields and aspects of causal reasoning.

Finally, Chapter 8 is on the normative notion of “proportionality” (introduced by Yablo 1992), which highlights the different levels and grains of causal claims and can be applied in cases where we have to select which are the right variables for a causal analysis. It draws on recent debates on levels and causal selection.

When it comes to the big picture, viz. the human face of causation, there are two pillars on which the book stands. The first is an evolutionary/developmental perspective, which underlies his “functional account of causation”. On this evolutionary assumption, most features of causal cognition are actually present *because*

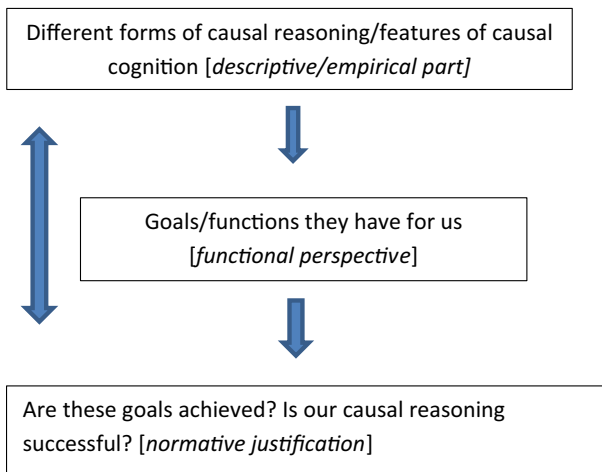
they contribute to some goals, selected “from among various possible forms of reasoning” as “those that best contribute to goals that we have” (29), where the “selective process might involve, among other possibilities, natural selection and/ or learning with feedback” (29). Hence, *causal cognition is* part of our evolutionary path as well as a larger part of our developmental processes, where the role of this part is even more emphasized, since its connection to the empirical literature is straightforward and can be empirically testable. For instance, in Chapter 3, Woodward draws on results from empirical studies in different developmental stages and across species in order to (i) obtain information on several aspects of human causal cognition and (ii) understand the reasons for its superiority from other primates in making successful causal judgments and predictions (see 158ff). He then uses these empirical studies as a means for extracting the normative notions of his theory.

The second pillar connected to the *human* face of causal cognition is the acknowledgement of our epistemic and calculational limitations. This is nicely shown when in Chapter 5 Woodward defends his choice of an “invariance-based” normative framework instead of accounts involving fundamental, exceptionless laws. He takes the hypothetical example of Larry, a Laplacian super-intelligence, who knows “everything”. Larry knows the exact causal structure of the world; therefore, he knows how everything fits together from first principles up to higher-level social phenomena. Larry would indeed not need “invariance” as a normative notion for measuring the strength of a causal relation or as a tool for discovering new generalizations or in order to predict what will happen at the next moment. However, given that *we* are humans with partial knowledge and not creatures with a Laplacian intelligence, “invariance” as a normative causal notion better captures our epistemic and calculational limitations. As he writes,

invariance-based notions seem to fit better with how we learn about and represent causal relations than accounts that invoke exceptionless fundamental laws—we start with only locally invariant relationships and then, in some cases, generalize and refine them into more and more invariant relationships, rather than thinking in terms of exceptionless laws from the outset. Exceptionless laws, if discoverable at all, are thus an extreme or limiting case of invariant generalizations, with the latter being the more generic notion. (260)

As we have noted, Woodward digs his favored “normative notions”, invariance and proportionality, out of empirical work on human causal cognition, thereby situating the source of these formalized norms into the descriptive/empirical. This step is a double-edged sword for Woodward’s account, since it takes the risk of his normative notions being entirely dependent on the descriptive features, thereby depriving them of their normative status. The grounding or justification of our evaluating norms in his account is a point that needs some special attention.

Indeed, there is a danger that the two-way street methodology of *CHF* might generate problems. We can depict the relations between the descriptive and the normative within the functional approach to causation, thus:



A potential problem here might be the following: the validation of the normative via the empirical requires a very careful methodology, a lot of testing, a lot of comparisons with other normative considerations. Indeed, Woodward repeatedly stresses the point that the descriptive does not necessarily align with the normative. For instance, he notes that "the fact that people conform to some pattern of judgment or reasoning does not by itself show that they are normatively correct to do so" (43). The required alignment relies on substantive assumptions such as "many people are 'rational,' in the sense that they make normatively appropriate causal inferences and judgments much of the time" (5). What saves this from being ungrounded in the sense of its normative part not being independent enough from the descriptive as to play its normative role is Woodward's initial powerful assumption that human cognition is useful in achieving certain goals. Thus, any normative considerations ought to play the role of assessing the relative success/failure of several features of our causal cognition in achieving these goals. At the end of the day, being a naturalist, Woodward takes it that all justification for the normative correctness of causal judgments is "a means/ends justification" (302). Once more, the functional perspective is Woodward's guide for extracting normative criteria from the descriptive work on causal cognition.

Funding Open access funding provided by HEAL-Link Greece.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission

directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Reference

Yablo, S. 1992. Mental Causation. *Philosophical Review* 101: 245–280.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.