



Click here to view
current issues
on the Chicago Journals website.

Review

Reviewed Work(s): *Essays on Explanation and Understanding* by J. Manninen and R. Tuomela

Review by: David Papineau

Source: *The British Journal for the Philosophy of Science*, Vol. 29, No. 3 (Sep., 1978), pp. 304-310

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/687026>

Accessed: 11-05-2021 07:58 UTC

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

JSTOR

greater simplicity, but not because it has greater content, for it does not—both formulae make equally precise predictions. This sort of simplicity is rather a matter of paucity of variables, and simplicity of kind of mathematical relationship. But there is no reason to regard this simplicity as mediated *via* the simplicity of the type of equation to which each belongs (*e.g.* the second equation as being a seventh-degree polynomial), and the latter as dependent on the “sample coverage” of that kind of equation. For low sample coverage may make for high support, but it does not make for high probability. Simplicity is not just a matter of lack of content, and Rosenkrantz has got in a bad muddle by supposing that it is.

If Rosenkrantz had taken more examples from physics and less typical statistician’s examples, he might have seen this. For it is typical of the kind of hypothesis with which statisticians deal that they do not differ greatly from each other in simplicity of the kind which I have just delineated. “ $p = 0.43$ ” is hardly more or less simple in this respect than “ $p = 0.4435$ ” or “ $0.42 \leq p \leq 0.44$ ”; although the last has much greater prior probability than the first two because of its less precise character. Physicists however often construct hypotheses (and could easily construct a very large number of hypotheses) equally successful in predicting observations so far, but differing in the number and kind of variables postulated, and the mathematical relations holding between them. Most of these they would never seriously think of putting forward, but the reason is that these are not simple enough in my sense.

R. G. SWINBURNE
University of Keele

MANNINEN, J. and TUOMELA, R. (eds.) [1976]: *Essays on Explanation and Understanding*. Dordrecht: D. Reidel. Pp. 440.

In his [1971] book G. H. von Wright was primarily concerned to defend two theses. The first was that the explanation of an agent’s actions in terms of his intentions conforms to a distinctive non-causal pattern of explanation (teleological explanation). The second was that the notion of cause itself depends conceptually on the idea of human action. A subsidiary theme of the book was that the human sciences, and in particular history, are ‘quasi-causal’ disciplines, in which teleological understanding plays an essential part in linking sequences of natural causation.

Von Wright argued that the basic procedure for explaining an action was (roughly) the reconstruction of the following kind of practical inference:

A intends to bring about *p*
A considers that he cannot bring about *p* unless he does *a*

Therefore, *A* sets himself to do *a*.

He denied that this was a schema of causal explanation on the grounds that the explanans is not linked to the explanandum by any empirical generalisation. On the contrary, he took the link to be conceptual. The premisses of the inference *entail* the conclusion, and, correlatively, cannot be established independently of our establishing the conclusion.

Causal explanation itself depends on the distinction between nomic generalisations, which support counterfactuals, and accidental generalisations, which do not. Von Wright argued that this distinction can only be understood in terms of a prior understanding of the possibility of human intervention in nature: one event is nomically antecedent to another just in case our having acted to bring about the former, on an occasion when it was in fact absent, would have eventuated in the occurrence of the latter.

Essays on Explanation and Understanding consists of fourteen papers addressed to the issues raised by von Wright's book, together with replies to these papers by von Wright himself, and a new essay 'Determinism and the Study of Man' in which von Wright somewhat modifies his position.

The first three papers are concerned with the broader issues of the status and subject matter of the human sciences. Manfred Riedel applauds the removal of history from the natural realm, and relates this to Hegel's teleological conception of historical development. But he suggests that Hegel's teleology is overly 'technical', explaining only the means and not the end, and recommends a return to Kant and the transcendental critique of historical methodologies. Riedel seems here to be sympathetic to 'critical theory', the contemporary German synthesis of Hegelian neo-Marxism with theories of the interpretative understanding of human action. This affiliation is more explicit in the paper by Rüdiger Bubner. Following recent writers like Habermas, Apel and Gadamer, Bubner suggests that a 'transcendental hermeneutics' can function simultaneously as a means for the interpretative understanding of human history and as an epistemological metatheory which comes to terms with the fact that all knowledge is historically conditioned. Maria Makai criticises von Wright from a more orthodox Marxist perspective. She objects to his dualism about man and nature. In her view a complete understanding of human action in its social and material context would transcend the traditional mind-matter dichotomy.

These first three papers are somewhat dense and likely to prove obscure to readers reared in an analytic environment. It would be unfortunate if they were thereby discouraged from continuing, as there is much that is worthwhile and accessible in the remaining contributions.

Most of the remaining papers focus on von Wright's scheme of intentionalist explanation. Jaegwon Kim discusses an apparent counter-example to this scheme originally raised by Nick Sturgeon: Jones is going to read a philosophy paper at the University of Cincinatti, and, since he is going to Cincinatti, he plans to look up his in-laws there. Now,

Jones intends to read a paper at the University of Cincinatti, and

Jones believes he cannot do this unless he goes to Cincinatti

Together these seem to explain quite adequately why

Jones sets himself to go to Cincinatti

But it is equally true that

Jones intends to visit his in-laws, and

Jones believes he cannot do this unless he goes to Cincinatti

Yet these facts do *not* seem to explain why

Jones sets himself to go to Cincinatti.

Von Wright in his reply to Kim tries to deal with the Sturgeon counter-example by means of a distinction between conditional and unconditional

intentions. Someone who unconditionally intends something is committed to doing whatever is required if he is to do that thing. Having a conditional intention to do something if and when certain conditions obtain does not, on the other hand, commit the intender to doing anything before those conditions are satisfied; in particular it in no way commits him to bringing about those conditions. Only unconditional intentions are qualified to explain the actions believed necessary to their fulfilment.

So the idea is that Jones only conditionally intends to see-his-in-laws-if-and-when-he-is-in-Cincinnati. Consequently this intention cannot feature in the explanans of a teleological explanation. The intention to read his paper, on the other hand, is unconditional, and so can explain his doing what is required to ensure this.

This response to the Sturgeon counterexample is inadequate. For Jones's intention to see his in-laws does explain quite happily his doing some of the things he believes necessary thereto, such as getting their phone number from his wife before he leaves. If von Wright is correct to distinguish between conditional and unconditional intentions, and to deny any explanatory power to the former, then Jones must presently intend unconditionally to see his in-laws in Cincinnati.

Of course there is still something dodgy about Jones's intention to see his in-laws, by comparison with his intention to read his paper. Jones's going to Cincinnati is essential to the fulfilment of either of these intentions, yet only the intention to read his paper seems to give an explanation of why he does so. But this is simply the asymmetry that Sturgeon raises, and which von Wright needs to give some account of.

Kim attempts to save von Wright's schema by attending to the way in which intentions are generated. He thinks that the way an intention stems from an agent's beliefs and wants is important for its explanatory significance. To deal with the Sturgeon case Kim proposes that if a given intention results from a belief that *p*, then that intention cannot explain the agent's intending to or setting himself to bring about that *p*. Jones's intention to see his in-laws, unlike his intention to read his paper, *results* from his belief that he is going to Cincinnati, and so cannot explain his going there.

In reaching this conclusion Kim is hinting at an alternative line of approach to action explanation. If the explanatory power of an intention depends on the way it issues from the agent's beliefs and wants, why bring in intentions at all? Why not simply explain actions as issuing directly from the agent's beliefs and wants? The distance between such an approach and the one von Wright actually has in mind is shown in his replies to those several contributors who do take him to be discussing the way actions issue from beliefs and desires, and criticise him accordingly. Somewhat surprisingly, von Wright does not resist their arguments. He merely observes that they are concerned with processes of deliberation and decision, whereas his problem was solely the further link between the formation of an intention and the ensuing action.

It is doubtful whether von Wright can legitimately restrict his brief in this way. In the end what the Sturgeon counterexample seems to show is that the route to the explanation of actions does not proceed *via* intentions at all, but comes directly from beliefs and wants. Raimo Tuomela suggests at the end of his

contribution that the Sturgeon case will cease to be a problem once we embrace the idea that actions depend causally on beliefs and desires. He does not spell this out in any detail. But (without worrying too much about the rather unfashionable causal aspect of his suggestion) it is easy enough to see how it would work.

Our understanding of the way actions depend on beliefs and desires tells us that Jones's going to Cincinnati depends on his desire to read his paper there, but not on his desire to see his in-laws. For, if he lacked any desire to read his paper there, he would not go. But, by contrast, he would still be going even if he had no desire to see his in-laws. This asymmetry does not stand if we replace 'desire' by 'intention'. It is easy enough to account for the asymmetry Sturgeon raises once we take beliefs and wants, rather than intentions, to be the primary explainers of actions.

From this point of view intentions as such do not explain actions. To say a man intends something is to say he desires that thing, either as an end or as a means, and that he is going to perform actions which he believes will further the satisfaction of that desire. The significance of the Sturgeon case is that it shows that those actions will not always depend counterfactually on that desire (even if this is so in most cases, which of course creates the impression that there is a direct explanatory link between intention and action). No doubt it would be possible to amend von Wright's scheme, *à la* Kim's modification, to restrict it to those intentions which 'point to' attitudes on which the relevant actions do depend. But this would be a quite *ad hoc* way of avoiding the conclusion that the explanation of action need not bring in intentions at all.

Let us now turn to von Wright's 'interventionist' analysis of causation. The most incisive comments on this are made by Peter Winch. Winch reconstructs von Wright's standard procedure for discovering causes as follows. We have a system in state α , which we assume will not change to a , nor subsequently display p , unless we change it to a . We change it to a , and p follows. This entitles us to the conclusion that a causes p . For we have established the relevant counterfactual that p wouldn't have been if a hadn't been. But this counterfactual presupposes the further counterfactual, that if we hadn't acted, a would have been absent. And this latter counterfactual is not a causal one, as our acting does not *cause* what we do. So causal judgments depend upon an *a priori* understanding of human actions.

Winch observes a number of faults in this. First, the crucial original assumption, that α won't change to a or p unless we intervene, itself seems to be a causal judgment. And this judgment seems independent of any of our ideas about human intervention. The fact we can intervene to produce a is irrelevant, since what we are after is what happens when we don't intervene, and if we are going to find this out from anywhere we will have to find it by just observing. Secondly, even if we allow, with von Wright, that our acting doesn't cause what we do, it by no means obviously follows that the connection between our *not* acting and what then happens is not a causal connection. So it is not obvious that 'If we hadn't acted, a would have been absent' is not a causal counterfactual. And, finally, it is quite unclear in exactly what sense this counterfactual is presupposed by the target causal counterfactual, that p wouldn't have happened if a hadn't.

Von Wright replies to the second of these criticisms by arguing that while causal counterfactuals state what alteration would have been produced if some change had occurred, 'action' counterfactuals state how things would have remained the same had we not done anything. This does not seem an adequate basis for the distinction he needs. Von Wright does not reply to Winch's other criticisms.

Of the other papers in the volume, those by Alasdair MacIntyre, Lars Hertzberg, Frederick Stoutland and Rex Martin discuss further the topics already mentioned. The rest are on topics only distantly related to von Wright's book.

In a paper entitled 'An Analysis of Causality' Aleksandar Kron develops some ideas originally put forward by H. A. Simon. Simon has shown how it is possible to define an asymmetrical cause-effect relation amongst the variables in certain solvable sets of linear equations. His approach explicates our intuitive notions about which variables are dependent on which. (Roughly, in the kind of system in question we can partially order the variables according to which other variables they appear with in equations. A variable which occurs in an equation which contains no other variables is independent of all others. A variable which occurs in an equation in which the only other variable is such an 'absolutely independent' variable, depends on the latter but is independent of all others. And so on.)

What Kron does is apply Simon's methods to the set of free variables in certain systems of well-formed formulae in the language of first-order logic. While this makes for a somewhat forbiddingly technical paper, it is of interest in suggesting that Simon's ideas might be of more general relevance than indicated in his original paper. On the other hand, it is most doubtful whether this should qualify as 'an analysis of causality'. Simon's techniques are intended to explicate the direction of causation amongst a set of variables that are already assumed to be nomically related. This is indeed an important problem in the theory of causation (and, incidentally, as Tuomela points out, one that von Wright's interventionist account completely fails to come to grips with). But, as Simon makes clear in his original paper (Simon [1953]), our ability to specify the direction of causation amongst the variables in some system of equations has to rest on an independent understanding of which variables depend on which. This is to do with the simple point that a given system of solvable equations can be reorganised by elementary row transformations without changing the solutions. These alternative ways of writing the equations will, however, present different relationships of relative dependence and independence amongst the variables in question. Simon argues that our choice amongst the various equivalent ways of writing the equations reflects our prior judgments about which variables would be affected and which unaffected were one or more of the non-zero coefficients in a given equation in the system to be altered. Clearly this means that Simon's work fails to yield a philosophical explanation of the direction of causation; nor does Kron present any reason to think his generalisation of Simon's work is an improvement in this respect.

The real interest of Simon's work lies in the connection it uncovers between causal ordering and 'identifiability'. A structure of equations is identifiable if we can identify certain fixed coefficients from a number of observations of the values of the variables on particular occasions. What Simon shows is that the

identifiability of a structure of equations requires that its variables be causally ordered; that is, that we know beforehand that the equations will represent a certain order of relative dependence and independence amongst the variables. This is of particular interest in connection with stochastic equations and the estimation of their coefficients from statistical data: work such as Simon's has been valuable in making clear exactly how much has to be assumed beforehand if we are to make meaningful use of multiple regression techniques. Kron says nothing about whether similar considerations arise with his generalisation of Simon's work. It is not obvious to me that they do.

Ilkka Niiniluoto is concerned to defend 'inductive-statistical' explanations against those critics who hold that such 'explanations' can only show that an event was to be expected, not *why* it actually occurred. To this end Niiniluoto follows Wesley Salmon in arguing that genuine inductive explanations are appropriate precisely to those events which are not completely determined. Such explanations are then inductive, not because we are ignorant of the factors that determine the explanandum, but simply because the full set of factors relevant to its occurrence fails to determine it. Niiniluoto then argues that this inductive pattern of explanation is appropriate to the explanation of historical events and of human actions in general. The difficulty with this is that in the kind of social scientific case Niiniluoto has in mind it is generally absolutely obvious that our explanations are incomplete—that the inclusion of (for practical reasons) unknown factors would alter the probability of the explanandum. The moral of this is not, I think, that such inductive explanations do not explain—often they patently do. Rather it is that Niiniluoto and Salmon are wrong to think that inductive explanation requires indeterminism. It would be much closer to the mark to see an inductive explanation as satisfactory precisely insofar as it succeeds in picking out some part of a set of factors which fully determines its explanandum.

Anthony Kenny contributes a precise and sophisticated analysis of the logic of the 'can' of human ability. He shows that this cannot be represented as a possibility operator in any of the standard modal systems, since it fails to distribute over disjunction. A standard person can either pick a red card or a black card from a standard pack; but it does not follow that either he can pick a red card or he can pick a black card. In the end Kenny suggests that the impossibility of representing this 'can' as a standard modality lends support to a conditional analysis of human abilities. Though Kenny does not go further, it is significant that such a conditional analysis would flow naturally from the view that actions depend on beliefs and desires: 'A can Φ ' could in the first instance be read simply as 'A will Φ , if his beliefs and desires indicate Φ -ing as the optimal thing to do'.

Jaako Hintikka examines the notion of intentionality. He rejects the traditional phenomenological view that what makes an act intentional is its being directed towards some content, and instead defends the equation of intentionality with intensionality, which he explicates as involving the simultaneous consideration of several possible worlds. Hintikka has interesting comments on Husserl's ideas and on possible world semantics. How seriously one takes the overall thesis depends on how plausible one finds his possible worlds. I wondered whether these were supposed somehow to underlie socially objective languages, or whether they were constituted by the subjective conjectures of individuals.

Either line raises problems: the former kind of possible worlds would seem to be unwarranted abstractions, while the latter kind would derive from certain of an individual's intentional acts, and so, as J. N. Findlay points out in a comment on Hintikka's paper, be ineligible for explaining away intentionality.

In 'Determinism and the Study of Man' von Wright refines his views on the intentional explanation of action; and he expands on the connection between such explanations and social phenomena like communication, norms and institutions. The latter observations are interesting enough. But there is nothing in the former remarks to remove the suspicion that it is a mistake to try to explain actions in terms of intentions.

DAVID PAPINEAU
University of Reading
Macquarie University

REFERENCES

- WRIGHT, G. H. VON [1971]: *Explanation and Understanding*. London: Routledge and Kegan Paul.
- SIMON, H. A. [1953]: 'Causal ordering and identifiability' in W. C. Hood and T. C. Koopmans (eds.): *Studies in Econometric Method*. New York: John Wiley.