

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

Editor: G. M. K. HUNT

assisted by V. GLADMAN

Review Editor: L. J. BENNETT

Subscriptions

Published at an annual subscription of £35 UK and Europe, US\$72 USA and rest of world; single issues at £11 UK and Europe, US\$23 USA and rest of world. Reduced rates are available for members of the British Society for the Philosophy of Science £17.50 UK and Europe, US\$35 USA and rest of world. Prices include postage by surface mail or, for subscribers in the USA and Canada by Airfreight or in Japan, Australia, New Zealand and India by Air Speeded Post. Airmail rates are available on request. Back issues are sold at the current single issue rate and complete back volumes are sold at the current single issue rate multiplied by the number of issues in the volume. The current plus two back volumes are available from the Oxford University Press, Walton Street, Oxford OX2 6DP. Previous volumes can be obtained from Dawsons Back Issues, Cannon House, Park Farm Road, Folkestone, Kent CT19 5EE, tel 0303 850101, fax 0303 850440.

Payment is required with all orders and subscriptions are accepted and entered by the volume/s. Payment may be made by the following methods:

Cheque (made payable to Oxford University Press)

National Girobank (Account 500 1056)

Credit Card (Access, Visa, American Express, Diners Club)

UNESCO Coupons

Bankers: Barclays Bank Plc, PO Box 333, Oxford, UK. Code 20-65-18. Account 00715654.

Please send orders and requests for sample copies to:

Journals Subscriptions Department, Oxford University Press, Walton Street, Oxford OX2 6DP.
Telefax (0865) 267773.

USA Distributor

The British Journal for the Philosophy of Science is published quarterly by Oxford University Press. Subscription is \$72 per year. Second Class paid at Newark, New Jersey. ISSN 0007 0882.

POSTMASTER: send address corrections to The British Journal for the Philosophy of Science, c/o Virgin Mailing and Distribution, Building 150, Newark International Airport, Newark, New Jersey 07114, USA.

Change of Editor

G. M. K. Hunt is retiring as editor of the journal after eight years in the post. From 1 October 1993 the journal will be edited by David Papineau. The new books editor will be Scott Sturgeon. From that date all typescripts and correspondence should be sent to the new editor at the Department of Philosophy, King's College London, Strand, London WC2R 2LS. E-mail: bjp@uk.ac.kcl.cc.oak. Facsimile: 071 873 2351. Books for review should be sent to the books editor at the editorial address.

© Oxford University Press 1993

All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without either the prior written permission of the Publishers, or a licence permitting restricted copying issued in the UK by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1P 9HE, or in the USA by the Copyright Clearance Center, 27 Congress Street, Salem, Mass 01970.

EDITORIAL BOARD

P. Alexander	Sir John Eccles	S. Körner	H. R. Post
G. Buchdahl	J. S. Giedymin	D. H. Mellor	M. L. G. Redhead
L. J. Cohen	M. B. Hesse	D. J. O'Connor	J. W. N. Watkins
A. Crombie	C. W. Kilmister	Sir Karl Popper	G. J. Whitrow

EDITORIAL ADVISORY PANEL

The Editor is helped by members of the Editorial Board, members of the Committee of the British Society for the Philosophy of Science and by an Advisory Panel which includes:

Jon Dorling	Barry Gower	Ernan McMullin	Paul Teller
Allan Franklin	Ian Hacking	David Miller	Peter Urbach
Grant Gillett	Colin Howson	Alan Musgrave	John Worrall
Donald Gillies	Henry Krips	Anthony O'Hear	Elie Zahar

Brit. J. Phil. Sci. 44 (1993), 187-203 Printed in Great Britain

Causation as a Secondary Quality

PETER MENZIES AND HUW PRICE

- 1 Introduction
- 2 The Agency Theory: A Probabilistic Version Outlined
- 3 Metaphysics Confused with Epistemology?
- 4 Unavoidable Circularity?
- 5 Unmanipulable Causes?
- 6 Unacceptable Anthropocentricity?

I INTRODUCTION

In this paper we defend the view that the ordinary notions of cause and effect have a direct and essential connection with our ability to intervene in the world as agents.¹ This is a well-known but rather unpopular philosophical approach to causation, often called the *manipulability* theory. In the interests of brevity and accuracy, we prefer to call it the *agency* theory.² Thus the central thesis of an agency account of causation is something like this: an event *A* is a cause of a distinct event *B* just in case bringing about the occurrence of *A* would be an effective means by which a free agent could bring about the occurrence of *B*.

In our view the unpopularity of the agency approach to causation may be traced to two factors. The first is a failure to appreciate certain distinctive advantages that this approach has over its various rivals. We have drawn attention to some of these advantages elsewhere, and we summarize below. However, the second and more important factor is the influence of a number of stock objections, objections that seem to have persuaded many philosophers that agency accounts face insuperable obstacles. In this paper we want to show that these objections have been vastly overrated.

¹ Among those who have proposed agency accounts of causation are Collingwood [1940], Gasking [1955], and von Wright [1975]. Perhaps Ramsey should also be regarded as an advocate of this view: in the later of his two papers on laws he says that 'from the situation when we are deliberating seems to me to arise the general difference of cause and effect' (Ramsey [1978], p. 146).

² It is not to be confused with the view of Richard Taylor that there is a special kind of causation by agents which is distinct from ordinary causation by events: see Richard Taylor [1966].

There are four main objections:

1. *Agency accounts confuse the epistemology of causation with its metaphysics.* It is widely conceded that experimentation is an invaluable source of evidence for causal claims; the objection is that it is a confusion to suppose that the notion of agency should thereby enter into the analysis of causal claims.
2. *Agency accounts are vitiated by circularity.* It is argued that the *bringing about* is itself a causal notion, and that this introduces a vicious circularity into an agency account.
3. *An agency account cannot make sense of causal relations between events which are outside the control of any agent.* For example, it is argued that such an account cannot make sense of the claim that the earth's revolution around the sun causes us to experience the seasons.
4. *Agency accounts make causation an unacceptably anthropocentric phenomenon.* Agency accounts are said to imply what is obviously false, namely that there would be no causal relations if there were no human agents (or different causal relations if there were different human agents).³

In responding to these arguments, our strategy will be to draw attention to the fact that there are closely parallel 'objections' to a quite uncontroversial feature of contemporary philosophical treatments of the concept of colour. Few philosophers would dispute that an adequate account of colour will need to make some reference to human perceptual states or capacities—that colour is a secondary quality, to use the familiar terminology. Of course, this common ground leaves plenty of room for disagreement, in particular as to precisely what form the required reference to human perception should take. As to the need for some reference, however, there is virtual unanimity. We have all learnt that much from Locke, if not already from Galileo, for example, or from Democritus. Our view is that the agency account of causation should be no more controversial.

Since we have neither the space nor the expertise to do justice to the full range of contemporary philosophical treatments of colour, we shall take as our reference point a simple version of the orthodox dispositional theory, namely the view that to be red is to be disposed to look red to a normal observer under standard conditions. This embodies the insight that colour is a secondary quality, defining the colour concept in terms of human capacities and responses. On the face of it, the reference to human capacities invites objections precisely analogous to the ones against the agency accounts of causation. Most of the objections have stock replies, however, and few philosophers would regard them as significant difficulties for the dispositional

³ For critiques of agency accounts of causation citing these objections, see Mackie [1976], Beauchamp and Rosenberg [1981], and Hausman [1986].

theory. Our argument is that analogues to these replies are available to defenders of the agency approach to causation. The parallel between the two kinds of theory is very close, and the agency account of causation should hence be no more controversial than the dispositional theory. In fact it should be less controversial, for the dispositional theory of colour is but one way of unpacking the insight that colour is a secondary quality. Our claim is simply that the agency theory correctly portrays causation as something analogous to a secondary quality—as a secondary quality, in fact, on a suitably extended understanding of that notion. Unlike the proponents of a dispositional theory of colour, however, we have no present need to take a stand on the internal politics of this general position.

The paper runs like this. In the next section, we sketch the probabilistic version of the agency account of causation, and outline some of its virtues (virtues we have each extolled at greater length elsewhere).⁴ In Sections 3–6 we then respond, in turn, to the above four objections to the agency approach to causation. In each case, as we have indicated, our defence is to exhibit the parallel between the argument in question and a strikingly unsuccessful objection to the dispositional theory of colour. We close by recommending a generalization of the usual notion of a secondary quality, a generalization which takes in causes as well as colours.

2 THE AGENCY THEORY: A PROBABILISTIC VERSION OUTLINED

As we have said, the common idea to agency accounts of causation is that an event *A* is a cause of a distinct event *B* just in case bringing about the occurrence of *A* would be an effective means by which a free agent could bring about the occurrence of *B*.⁵ The agency account we endorse develops this idea differently from the older agency accounts of Collingwood, Gasking, and von Wright. As we see it, these older agency accounts suffer from some significant limitations. In particular, they assume that causation is deterministic. In our view there is much to be said for the greater generality of an indeterministic approach. Such a notion of causality seems necessary in quantum mechanics, and arguably plays a major role in much more familiar contexts. And in any case, it seems reasonable to expect that an indeterministic theory will incorporate the deterministic notion, as a special or limiting case. Much of

⁴ See Menzies [forthcoming 1] and Price [1991a].

⁵ Strictly speaking our view is that the notion of agency must enter the account of *causal conditions* rather than of *causes*. It is a now familiar point that there are many more causal conditions for some effect than there are causes; and which causal conditions we choose to highlight as causes is a highly context-dependent affair. The customary treatment is to take the notion of causal condition as prior to that of cause; and to introduce a contextual variable to explain the principles by which we select the causes among the causal conditions. We endorse a treatment along these lines. But it simplifies matters to talk of causes rather than causal conditions, leaving unstated the principles for the selection of causes.

what we say here could be recast in terms of the special case; but we see no reason not to go directly for the general account.

Our approach invokes the means-end relation, and we characterize this in terms of what we call *agent probabilities*. Agent probabilities are to be thought of as conditional probabilities, assessed from an agent's perspective under the supposition that the antecedent condition is realized *ab initio*, as a free act of the agent concerned. Thus the agent probability that one should ascribe to *B* conditional on *A* (which we symbolize as ' $P_A(B)$ '), is the probability that *B* would hold were one to choose to realize *A*.

Another way to put this is that the agent probability of *B* conditional on *A* is the probability that should enter into the calculations of a rational agent whose abilities consist in the capacity to realize or to prevent *A*, and whose goals turn entirely on *B*. This characterization also marks a difference between us, concerning the nature of rational decision theory. One of us (HP) thinks that an evidential theory will suffice, and hence that agent probabilities may be thought of as a species of evidential probability. The other (PCM) believes that a non-evidential or 'causal' decision theory is required, and hence that objective conditional chances must enter into the formulation of agent probability.⁶ This difference is not germane to our present purposes. What is common to our views is that agent probabilities are to be functionally defined in terms of their role in rational decision-making; and that a minimal constraint on a correct account of rational decision-making is that it yields prescriptions for Newcomb problems which at least accord with those of causal decision theory.

Given that agent probabilities are thus construed in terms of their role in decision theory, they embody the basis of a formal analysis of the means-end relation. To say that *A* is an effective means by which an agent could achieve *B* is to say that if the agent were to have an overriding desire that *B* should obtain, then an adequate rational decision theory would prescribe realizing *A* rather than $\neg A$. (An agent has an overriding desire that *B* should obtain in a given decision situation just in case his or her preferences are such that *B* is more desirable than $\neg B$, whichever option is realized.) But it is easily shown that the expected utility of *A* is greater than $\neg A$ in these circumstances if and only if $P_A(B)$ is greater than $P_{\neg A}(B)$.⁷ In other words, *A* constitutes a means for achieving *B* just in case $P_A(B)$ is greater than $P_{\neg A}(B)$.

Of course, this condition expresses the central idea of probabilistic theories of causation, namely that the difference a cause makes to its effect is that its occurrence increases the probability of the effect's occurrence.⁸ Thus our

⁶ For the first view, see Price [1991a] and [1986]; for the second view, see Menzies [forthcoming 2].

⁷ The argument is given by Mellor [1988]. We are generally indebted to this article for its insights into how an agency account of causation can be spelled out in probabilistic terms.

⁸ The first theory to spell out this idea with any precision was that of Suppes [1970]. The central idea of probabilistic theories applies only to straightforward instances of causation that do not involve pre-emption or overdetermination. One way to develop the central idea to deal with these more complex cases of causation is proposed in Menzies [1989].

agency theory of causation is *inter alia* a probabilistic theory of causation. What makes it different from, and in our opinion superior to, other probabilistic theories is precisely its appeal to the notion of agency, by way of the notion of agent probability.

We have argued elsewhere that this use of agent probabilities enables the present account to circumvent several difficulties facing orthodox probabilistic theories.⁹ In particular, the account avoids the problem of 'spurious' evidential dependencies of causes on effects, and between the joint effects of a common cause.¹⁰ Briefly, this is because spurious correlations reflect the evidential connections between events that arise as an indirect reflection of the existence of genuine causal relations. For example, the evidential relation between an effect and its cause indirectly reflects the causal relation running from the cause to the effect; and the evidential relation between the joint effects of a common cause reflects the genuine causal relations running from the common cause to its two effects. Agent probabilities fail to generate spurious correlations because they abstract away from the evidential import of an event, in effect by creating for it an independent causal history. For example, in enquiring whether one's manipulation of an effect *B* would affect the probability of its normal cause *A*, one imagines a new history for *B*, a history that would ultimately originate in one's own decision, freely made.¹¹ And one thereby deprives *B* of its usual evidential bearing on *A*. The same is true, *mutatis mutandi*, for the situation in which one manipulates one effect of a common cause to see whether it makes a difference to the probability of another effect.

We take the fact that the virtues of agent probabilities have been overlooked by orthodox probabilistic theories to be an unfortunate legacy of the observation-oriented empiricism which Hume bequeathed to subsequent discussions of causation. Given this orientation, the Humean tradition takes for granted that causation is to be analysed in observational terms. Thus it is no surprise that Humean analyses give pride of place to constant conjunctions or relative frequencies. In our view the lesson of spurious probabilities is that the Humean strictures are too severe. Empiricists need to keep in mind that human subjects have access to the world in two ways: as observers, certainly, but also as *agents*, capable of intervening in the processes of the world at will.

As we have indicated, we take it that one of the advantages of this extended empiricism is the ability to deal with the worst of the problems afflicting orthodox probabilistic theories of causation. The main task of this paper is to exhibit another advantage. We want to show that in admitting action on a par

⁹ See Price [1991a], Menzies [forthcoming 1].

¹⁰ *B* probabilistically depends on *A* if and only if $P(B/A) > P(B/\neg A)$, or equivalently $P(B/A) > P(B)$. These relations are symmetric if the probabilities are standard conditional probabilities. Spurious dependency is discussed by Patrick Suppes in [1970], Chapter 2. See also Cartwright [1979].

¹¹ 'In a sense my present action is an ultimate and the only ultimate contingency', as Ramsey ([1978], p. 146) puts it.

with perception as a means of access to the world, it also has the resources to meet the various objections that have seemed to show that agency cannot, in principle, play a constitutive role in an account of causation. It is to these objections that we now turn. As promised, our main strategy will be to show that to each objection to the proposed use of agency in an account of causation there is a parallel and no less forceful 'objection' to an uncontroversial use of visual capacity in an account of colour.

3 METAPHYSICS CONFUSED WITH EPISTEMOLOGY?

Doesn't the agency approach to causation confuse the epistemology of causation with its conceptual analysis? No doubt experimentation has a privileged position in the empirical methodology for testing causal claims. But surely this is no reason for supposing that the notion of agency has a constitutive role in the analysis of causation: after all, it is the cardinal sin of verificationism to suppose that the means by which a statement is verified or tested determines the meaning of the statement.

In response to this objection, we first protest that verificationism is no part of our brief: we disavow any view which states that the meaning of causal statements is to be given in terms of their verification-conditions or the conditions for their warranted assertion. In our view, the notion of agency enters the analysis of causation not because experimental manipulations provide evidence for causal claims, but because the notion of causation is an extrinsic one which by its very nature is rooted in the idea of manipulation.

To explain our point, let us turn to the colour analogy. For definiteness we have settled on the dispositional theory of colour, according to which an object is red, say, just in case it would look red to a normal observer under standard conditions.¹² This theory makes colour a secondary quality in the sense that the concept of colour is taken to be an extrinsic or relational one, where the constitutive relation is to a certain kind of human response: in the case of the colour red, the 'looks red' response. Clearly, such a theory has epistemological implications, since it has as a consequence that the best kind of evidence for an object's actually being red is that it should look that way to a normal observer under the right conditions. But it would be a misunderstanding of the nature of this kind of theory to suppose that it erroneously introduces epistemological considerations into the analysis of colour. The dispositional theory does not

¹² We emphasize that in comparing the agency approach to causation with the dispositional theory of colour, we do not wish to commit ourselves to the latter as a correct account of colour. As we noted earlier, any account of colour which makes the notion of colour at least conceptually dependent on human capacities and responses would have sufficed for our purposes. Even a physicalist account would have done: for one such theory see Jackson and Pargetter [1987]. For some reasons against taking the dispositional strategy to provide the best generalization of the notion of a secondary quality, see Price [1991b].

depend on any contentious verificationist assumption, but merely on the assumption that colour is essentially a secondary quality.

We propose that the agency approach to causation should be seen in the same light, so to speak. This approach makes the notion of causation an extrinsic one, to be explained by relation to our experience as agents: on our probabilistic version of the agency approach, the concept of causation is to be explained in terms of the way in which an agent's producing, manipulating, or 'wiggling' one event affects the probability of another event. It is apparent that this analysis has a bearing on epistemological issues, in that it implies that the best kind of evidence for a causal claim comes from experimental manipulations. But once more it would be a serious mistake to suppose that this account depends on a verificationist fallacy. Rather the central point is that the concept of causation is to be explained by relation to our experience as agents in the same way that the concept of colour as a secondary quality is to be explained by relation to our experience as observers.

Of course, the fact that our experiences as observers and as agents play such a central role in the explanation of colours and causation does not mean that these experiences are infallible guides to the existence of these properties and relations. For example, it is a familiar point that something can look red to someone although it is not actually red because it is observed under non-standard conditions. The dispositional theory of colour clearly makes allowance for this kind of discrepancy. We find similar discrepancies between our experiences as agents and causal relations: for example, the relative frequencies discovered through experimentation may point to a certain causal hypothesis which is in fact false, the frequencies deviating by chance from the true mark; to put it another way, actions may succeed by accident. An agency theory of causation, if properly formulated, should allow for this possibility. This is certainly the case with our preferred probabilistic version, according to which experimental relative frequencies provide defeasible evidence for the agent probabilities.

4 UNAVOIDABLE CIRCULARITY?

The second and perhaps most frequently cited objection to the agency theory is that it necessarily involves a vicious circularity. The apparent circularity is plain enough in our informal statement of the agency approach, according to which *A* is a cause of a distinct event *B* just in case *bringing about A* would be an effective means by which a free agent could *bring about B*. This statement contains two references to 'bringing about', which seems on the face of it to be a causal notion: doesn't an agent bring about some event just in case she causes it to occur? It would appear, then, that agency accounts are vitiated by the fact that they employ as part of their analyses the very concept which they are trying to analyse.

Once more we observe that exactly the same objection can be lodged against the dispositional theory of colour. As several philosophers have observed, the theory seems to be viciously circular in an obvious way: for in stating that to be red is to look red to a normal observer in standard conditions, the theory employs the concept 'red' on the right-hand side of the analysis.¹³ Our view is that the two kinds of theory stand or fall together on this point.

A possible defence against this kind of criticism is to argue that the theories in question are not meant to be reductive analyses which reduce the concepts of causation and colour to their atomic constituents. Circularity is admittedly a decisive flaw in an analysis, so conceived. But if all that these theories attempt to do is to state some interesting interrelationships between concepts, then the circularity objection is not decisive: for even a circular account of a concept can be informative as a statement of how the concept is intertwined with others, so long as it is not trivial. The dispositional theory of colour and the agency approach to causation could then be defended as non-trivial, albeit circular, accounts of their respective concepts.

While we have the sympathy with the spirit of this reply, we believe that a stronger defence is possible. We wish to argue that despite appearances to the contrary, these theories do not involve any circularity. As before, we start with the dispositional account of colour. The key to seeing that this theory is not circular is to recall that colour terms, like the terms for other secondary qualities, can be introduced by ostension. Thus a novice can be introduced to the concept 'looks red' by being shown samples of red: the salience of the redness in the samples and the novice's innate quality space should suffice for him to grasp the fact that the samples look alike in a certain respect. Thus, the dispositionalist explanation of the concept 'red' need not fall into the trap of circularity. The dispositionalist can explain the concept 'looks red' by ostensive definition, without having to rely on any colour concept.

A similar story may be told in the agency case. The basic premiss is that from an early age, we all have direct experience of acting as agents. That is, we have direct experience not merely of the Humean succession of events in the external world, but of a very special class of such successions: those in which the earlier event is an action of our own, performed in circumstances in which we both desire the later event, and believe that it is more probable given the act in question than it would be otherwise. To put it more simply, we all have direct personal experience of doing one thing and thence achieving another. We might say that the notion of causation thus arises not, as Hume has it, from our experience of mere *succession*; but rather from our experience of *success*: success in the ordinary business of achieving our ends by acting in one way rather than another. It is this common and commonplace experience that licenses what amounts to an ostensive definition of the notion of 'bringing

¹³ For a recent statement of this objection see Boghossian and Velleman [1989].

about'. In other words, these cases provide direct non-linguistic acquaintance with the concept of bringing about an event; acquaintance which does not depend on prior acquisition of any causal notion.¹⁴ An agency theory thus escapes the threat of circularity.

It is true that Hume himself considers the possibility that 'the idea of power or necessary connexion' derives from the fact that we are conscious of 'the internal power . . . of our will'. He rejects it on the grounds that in this internal domain, succession is no more evidence of necessary connexion than it is elsewhere. However, we want to emphasize that we are not appealing particularly to such inner experience. It may be, on our view, that agents are sometimes conscious of successfully achieving the motion of their bodies in the light of their intentions; but nothing depends on this. The basic cases on which we rely are those in which both means and end are non-psychological events.

5 UNMANIPULABLE CAUSES?

A third common objection to agency accounts is that they cannot explain causal relations between unmanipulable events. As an example of such causal relations, one might cite the causal claim that the 1989 San Francisco earthquake was caused by friction between continental plates. Obviously, no normal agent has the capacity to bring about friction between continental plates. The agency approach, then, appears to be seriously defective in not being able to accommodate such straightforward causal claims.

The essential feature of this kind of example is that it is physically impossible, given the capacities of a normal agent, to manipulate the cause and effect at will. It is easy to see what the parallel counterexamples are for the dispositional theory of colour: these are cases in which it is physically impossible for a normal observer to observe coloured things under standard conditions. One such example concerns the photosphere, which is the region inside the sun immediately surrounding its core.¹⁵ From the conjectured physical character of its molten contents, the photosphere is thought to emit light of the kind which would look red in other circumstances. This kind of example is a difficulty for the dispositional theory precisely because it is physically impossible for a normal observer to get into a position to observe the photosphere under standard conditions.

One might try to defend both theories against their respective counterexamples by arguing that, contrary to first appearances, the relevant

¹⁴ Of course, to say that one has immediate experience of causation through the exercise of the will falls short of Berkeley's view that all genuine instances of causation involve the exercise of the will. For Berkeley's view, see *The Principles of Knowledge*, Sections 30-2, 51-3, 60-6, 102-8.

¹⁵ We owe this example and several others in this section to Mark Johnston, who mentioned them as difficulties for the dispositional theory of colour in a seminar 'Colour as a Philosophical Paradigm' at the ANU in August 1989.

dispositional counterfactuals do hold in such cases. For it is a well-known feature of their application to deterministic processes that the relevant kind of counterfactual is well defined even though its antecedent describes a physically impossible situation.¹⁶ Thus, it might be argued that it is in fact true that if, *per impossibile*, an observer were to observe the photosphere, it would look red to him; and that it is true that if, *per impossibile*, an agent were able to manipulate continental plates, he would thereby be able to bring about earthquakes. Perhaps one might try to make these counterfactuals plausible by invoking a conception of an ideal observer or agent, a conception which abstracts away from the usual limitations of human perception and manipulation.

Whatever the intrinsic merits of this suggestion, we do not believe that it can be the whole story. For there are certain difficult examples which show that it is not possible to cash out a dispositional theory of colour or of causation in straightforward counterfactual terms. These difficulties point to a better way of handling the style of objection under consideration.

Let us consider the difficulties as they arise for the dispositional theory of colour. These difficulties highlight the fact that a thing can be coloured even though it would not look coloured to a normal observer under standard conditions. A good example illustrating this fact concerns the chemical substance rhodopsin, which changes its light-reflective properties when many photons impinge on it. In the dark rhodopsin has the molecular structure of something which reflects yellow-looking light. When it is exposed to light, however, it undergoes a chemical change whose effect is to make it reflect red-looking light. Here it is plausible to say that rhodopsin has the surface colour yellow, even though it does not look yellow to a normal observer under standard conditions. The surface-colour of rhodopsin is what has come to be called a 'finkish' disposition, that is a disposition which vanishes when it is put to the test.¹⁷

The other problematic kind of case has to do with dispositional properties which are masked by other properties. For example, material made of fluorescein is described as having the surface-colour green, even though it radiates a bright orange colour under standard lighting. The explanation is that the material in fact has two dispositions, the first to reflect green-looking light and the second to absorb ultraviolet light and then radiate orange-looking light. The effects of the latter disposition swamp those of the former. The fact that fluorescein is said to have the surface-colour green, even though it looks bright orange under standard lighting, obviously presents a difficulty for the dispositional theory of colour.

However, it is not difficult to see how the dispositional theory can be made

more sophisticated so as to accommodate such examples. For while fluorescein does not look green under standard lighting, it is reasonable to suppose that there is some object sharing some of its intrinsic properties which does look green under these conditions. In particular, those physical properties which underly its disposition to reflect green-looking light might be possessed by another kind of material lacking the additional masking properties. This other kind of material would be such as to make true the counterfactual to the effect that it would look green if a normal observer were to observe it under standard lighting. The cases of rhodopsin and of the photosphere can be handled in a similar way: rhodopsin or the photosphere have intrinsic properties which are very similar to, if not identical with, those of objects which make the appropriate dispositional counterfactuals true.

From the fact that fluorescein, rhodopsin, and the photosphere have intrinsic properties which resemble those of objects which make the appropriate dispositional counterfactuals true, a plausible principle of analogical reasoning permits us to draw the conclusion that fluorescein reflects green light, that rhodopsin is yellow, and that the photosphere is red. It is a simple matter to weaken the dispositional theory of colour to allow for the operation of this principle of analogical reasoning. So weakened, the theory allows that an object has a certain colour not only if it possesses intrinsic properties which make true the appropriate counterfactual but also if it possesses intrinsic properties which are identical with or closely similar to those of an object of which the appropriate counterfactual is true.

The important point is that we can handle the analogous objection to agency accounts of causation in the same way. For we would argue that when an agent can bring about one event as a means to bringing about another, this is true in virtue of certain basic intrinsic features of the situation involved, these features being essentially non-causal though not necessarily physical in character. Accordingly, when we are presented with another situation involving a pair of events which resembles the given situation with respect to its intrinsic features, we infer that the pair of events are causally related even though they may not be manipulable. Once more, this inference relies on the principle of analogical reasoning noted above. The agency account can be weakened to allow for the application of this principle in much the same way as the dispositional theory of colour was weakened above. In its weakened form, the agency account states that a pair of events are causally related just in case the situation involving them possesses intrinsic features that *either* support a means-end relation between the events as is, *or* are identical with (or closely similar to) those of another situation involving an analogous pair of means-end related events.

Clearly, the agency account, so weakened, allows us to make causal claims about unmanipulable events such as the claim that the 1989 San Francisco earthquake was caused by friction between continental plates. We can make

¹⁶ For a discussion of this point, see Lewis [1979].

¹⁷ David Lewis ([1989], p. 117) credits the term 'finkish dispositions' to Ian Hunt.

such causal claims because we believe that there is another situation that models the circumstances surrounding the earthquake in the essential respects and does support a means-end relation between an appropriate pair of events. The paradigm example of such a situation would be that created by seismologists in their artificial simulations of the movement of continental plates.

6 UNACCEPTABLE ANTHROPOCENTRICITY?

Finally, it is often felt that agency accounts make causation an unacceptably anthropocentric phenomenon. In its most naïve form, the objection is that if the agency approach were correct, it would follow that there could be no causal relations at times or places at which there are no agents.

This form of the objection is easily deflected. It is important to bear in mind that agency theories do not say that causal relations exist only when agents have *actually* performed the appropriate experimental manipulations, any more than dispositional theories of colour say that colours exist only when observers have *actually* experienced colour sensations. The point is that both kinds of theory, being dispositional in character are properly understood as having counterfactual breadth. Thus the dispositional theory of colour is to be understood as stating that an object is red just in case it is true that *if* a normal observer *were* present and were to observe the object under standard conditions, it *would* look red to her; and an agency theory of causation is to be understood as stating that a causal relation exists between two events just in case it is true that *if* a free agent *were* present and able, she *could* bring about the first event as a means to bringing about the second. (We have just seen that in both cases, if necessary, the counterfactuals concerned may be thought of as secured via lower-level theoretical similarities.)

However, there is a more sophisticated form in which the anthropocentricity objection may be pressed. It may be argued that the agency approach to causation implies, incorrectly, that in possible worlds in which agents have different powers from those they possess in the actual world, different causal relations obtain. For example, consider a possible world which is exactly like the actual world except that by an accident of natural selection agents have a much more restricted set of manipulative powers. At first sight, it might appear that the agency approach is committed to saying, quite implausibly, that the truncated manipulative powers of the agents imply a truncated set of causal relations.

The parallel objection in the case of colour is of course that the dispositional theory implies that in a possible world in which, say, normal observers are red-green colour-blind, there is no distinction between red and green things. The standard reply to this objection about colour is to rigidify the relevant dispositions, anchoring them to the perceptual capacities of the normal

observers of the *actual* world, not those of the possible world in question.¹⁸ It follows that red can still be distinguished from green in the possible world in which people are all red-green colour-blind; for an *actual* normal observer could still distinguish them, even though a normal observer of the other possible world could not. Clearly, this strategy is also available to a defender of the agency approach to causation. The point will be that the powers relevant to identifying the causal relations obtaining in any possible world are the powers of *actual* agents, rather than those of the *possible* agents of that world. Consequently, it is to be expected that the full range of causal facts would still hold in the possible world under consideration.

In both cases, however, it may be felt that the strategy of rigidifying the relevant dispositional facts does not go to the heart of the objection. In a related context, for instance, David Lewis has recently complained that the use of this strategy in defence of a dispositional theory of value does more to hinder the expression of the worry underlying the objection than to make it go away.¹⁹ He argues that it is possible to make the worry re-emerge. Rephrased so that it applies to the case of causation, his argument runs as follows. Suppose that the world had developed in such a way that we had fewer manipulative abilities and skills than we actually possess but that we still applied our concept of causation roughly in conformity with the agency approach. In this case, the reference of the expression 'relation between events such that an actual agent could manipulate one event as a means to bringing about the other' would have been fixed on different relations, even though our way of fixing the reference would have been the same. In one good sense, Lewis argues, we would have meant by 'causes' just what we actually mean but it would be true for us in these circumstances to say, of some events A and B, that A did not cause B, even though by our current lights A did cause B.

The point of Lewis's argument is thus that in virtue of their reference to human responses or capacities, the theories under consideration imply an untoward kind of relativity in their respective subject matters. Our response is to accept that this kind of relativity is a consequence of the theories concerned, but to deny that it is untoward. We make two main points in support of this conclusion. The first, as usual, is that the characteristic of causation thus identified is already a fairly non-problematic feature of colour and the other classical secondary qualities. It is something we live with in those cases, and may be expected to accommodate ourselves to in the case of causation. Secondly, however, we want to point out that there is an important difference of degree between the two cases. As we shall explain in a moment, it turns out that causation is very much less sensitive than colour, say, to the accidents of the human situation. In this we find a basis for the intuition that causation is

¹⁸ This reply is given, for instance, by Smith [1987].

¹⁹ Lewis [1989], especially pp. 132–3.

significantly more 'objective' than the usual secondary qualities—an intuition with which we thus concur.

Of the usual secondary qualities, it is perhaps those of taste that best exemplify the general point of Lewis's argument. Lemons are normally sour, for example. However, we humans might have evolved in such a way that our taste buds were less sensitive to sour things and more sensitive to sweet things. Vinegar would still have tasted sour, but lemons would then have seemed as sweet as (actual) ripe oranges. In such a world it would thus have been generally agreed that, unlike vinegar, lemons are *not* sour. Grant that lemons are the same in themselves in this counterfactual world as in the actual world, and that in the obvious sense 'sour' means the same in the both worlds, and we have Lewis's conclusion: sourness is in the eye of the beholder, so to speak.

In the case of taste this conclusion seems not only uncontroversial, but also unproblematic, at least in a pragmatic sense. That is, it is clear that even this quite prominent degree of potential relativity does not prevent speakers from employing such notions as sourness and sweetness in useful ways, in science as well as in ordinary life. By analogy, we are inclined to argue that even if the agency theory of causation were to imply that the truth about causation is relative to human capacities in the same way and to the same extent, that consequence should simply be counted a discovery about the notion of causation. Certainly it would not conflict with most ordinary use of the notion.

At the same time, however, we share the intuition that this is not how things will turn out. On the contrary, we think that the extension of the notion of causation is very much less sensitive to possible variations in the human condition than are notions of taste and colour. To exhibit the basis of this difference, let us go back to the non-trivial form of the anthropocentricity objection, namely the claim that the agency approach implies that in possible worlds in which agents have different powers from those they possess in the actual world, different causal relations will obtain. In the previous section we saw that by appealing to a principle of analogical reasoning an agency approach may extend its scope well beyond the domain of those things in a particular world that the agents of that world can *actually* influence. (This was the gist of our reply to the non-manipulability objection.) In consequence, it is far from clear that any modification of mere *degree* in our powers as agents will issue in any modification in the causal relations we are thus inclined to ascribe. On the contrary, it seems that agents with different capacities will nevertheless envisage the same range of *possible* causal relations, provided that they employ the principle of analogical reasoning we noted earlier as licensing the extrapolations of their manipulative capacities.

This suggests that in the case of agency, the only relevant possible world for the purposes of the anthropocentricity objection is the limiting case: the world in which, like Dummett's intelligent trees,²⁰ cognitive beings have *no* powers

²⁰ See Dummett [1978], p. 333.

as agents. Concerning this case, we accept that if there is no notion of agency to start with then 'nothing will come of nothing'—no amount of modelling and theoretical similarity will be of any use, if it has no material with which to start. In this case, however, we contend that such creatures could develop *no* notion of causation.²¹ In other words, we suspect that for causation, unlike for taste or colour, the capacity-dependent options are limited to the choice between a notion with the application it actually has, or no notion at all. So the agency approach does not conflict with the intuition that causation is significantly 'more objective' than colour or taste; on the contrary, it provides a plausible justification for this intuition. For it identifies a very significant difference of degree between causation on the one hand, and secondary qualities such as colour and taste on the other.

Of course, some readers may feel that a mere difference of degree just isn't good enough; that there must be some more principled distinction between the 'objectivity' of causation and the 'subjectivity' of the ordinary secondary qualities. We disagree, and take the view that the onus now lies with our opponents. We acknowledge, however, that the present discussion falls in the shadow of larger concerns. What is it for a property to be objective rather than subjective? Should we expect a sharp distinction anywhere, or could it all be a matter of degree?²² These issues are beyond the scope of this paper.²³ Our concern here is much more limited. We are simply trying to show that provided one is prepared to countenance the possibility that causation has the character of a secondary quality, the standard objections to the agency theory appear no more forceful than the analogous arguments against orthodox accounts of colour.

In our view the best way to characterize these parallels between causation and colour is to say that both can be viewed as secondary qualities under a generalized understanding of this notion. The usual characterization of a

²¹ In this we may seem to disagree with Dummett, who says that he is 'inclined to think that we could have some kind of concept of cause, although different from that we now have, even if we were mere observers and not agents at all—a kind of intelligent tree' (Dummett [1978], p. 333). The disagreement is largely terminological, however. Dummett argues that in a temporally asymmetric world such as we inhabit, even mere observers might find a need for an asymmetric mode of explanation. Certain observed events might be seen as explicable only in terms of *earlier* events, for example. But in the absence of a notion of causation grounded in agency, there seems little reason to call this kind of explanation 'causal'. After all, there is nothing except the accident of its temporal orientation to distinguish it from what we would normally think of as *non-causal* teleological explanation. Whatever we call it, however, it does not seem to be the full-blooded notion of causation with which we are here concerned.

²² One thought that might lie behind the above concern is that the notion of causation will itself have to play a crucial role in distinguishing those matters that are fully objective—those on which our judgements are simply *caused* by the world—from those that should have seen, in whole or in part, as 'coming from us', as projections of our responses to the world. See, for example, Blackburn [1984], Chapters 6 and 7 (particularly pp. 212 and 257).

²³ They are tackled from a rather different perspective in Price [1988], particularly Chapter 8; see also Price [1991b].

secondary quality, as a quality which tends to elicit a characteristic sensory experience in human subjects under specified conditions, is too restrictive in two ways. First, it confines attention to monadic properties, leaving no room for relations such as causation. Secondly, and more importantly, it applies only to those properties which have a sensory import. As such, it perpetuates a constant philosophical preoccupation with passive observation to the neglect of active intervention in the world: it should be kept in mind that we interact with the world not only as observers but also as agents. We advocate the adoption of a more general notion of a secondary property, which expressly disavows both these restrictions.²⁴ With such a notion in hand, one might then turn to the enticing question as to how much it includes. Causal relations as well as colours, we think, but how much else besides?

ACKNOWLEDGEMENT

We are grateful to Frank Jackson, Mark Johnston, Karen Neander, and Philip Pettit for many helpful discussions of this material.

Research School of Social Sciences
Australian National University
Canberra, Australia 2601
School of Philosophy
University of Sydney
Sydney, Australia 2006

REFERENCES

- BEAUCHAMP, T. and ROSENBERG, A. [1981] *Hume and the Problem of Causation*. Oxford University Press.
- BLACKBURN, S. [1984]: *Spreading the Word*. Oxford University Press.
- BOGHOSSIAN, P. and VELLEMAN, J. [1989]: 'Colour as a Secondary Quality', *Mind*, 98, pp. 81–103.
- CARTWRIGHT, N. [1979]: 'Causal Laws and Effective Strategies', *Noûs*, 13, pp. 419–37; reprinted as Chapter 1 in Cartwright [1983]: *How the Laws of Physics Lie*, Oxford University Press.
- COLLINGWOOD, R. G. [1940]: *An Essay in Metaphysics*, Oxford University Press.
- DUMMETT, M. [1978]: 'Bringing About the Past', in *Truth and Other Enigmas*, pp. 333–50. Duckworth.
- GASKING, D. [1955]: 'Causation and Recipes', *Mind*, 64, pp. 479–87.

²⁴ It is not yet clear what form such a generalization should take. One possible strategy is to take as one's model the familiar dispositional treatment of colour. Mark Johnston takes this route in characterizing what he calls *response-dependent* concepts: see his [1989] and [forthcoming], and also Pettit [1991]. For an alternative approach, criticizing the use of the notion of response-dependence in these contexts, see Price [1991b].

- HAUSMAN, D. [1986]: 'Causation and Experimentation', *American Philosophical Quarterly*, pp. 143–54.
- JACKSON, F. and PARGETTER, R. [1987]: 'An Objectivist's Guide to Subjectivism about Colour', *Revue Internationale de Philosophie*, 41, pp. 127–41.
- JOHNSTON, M. [1989]: 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society*, Supp. Vol., 63, pp. 139–74.
- JOHNSTON, M. [forthcoming]: 'Objectivity Refigured', in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*, Oxford University Press.
- LEWIS, D. [1979]: 'Counterfactual Dependence and Time's Arrow', *Noûs*, 13, pp. 455–76.
- LEWIS, D. [1989]: 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society*, Supp. Vol., 63, pp. 113–37.
- MACKIE, J. [1976]: 'Review of Causality and Determinism', *Journal of Philosophy*, 73, p. 216.
- MELLOR, D. H. [1988]: 'On Raising the Chances of Effects', in J. H. Fetzer (ed.), *Probability and Causation*, Kluwer, pp. 229–39.
- MENZIES, P. [1989]: 'Probabilistic Causation and Causal Processes', *Philosophy of Science*, 56, pp. 642–63.
- MENZIES, P. [forthcoming 1]: *Singular Causation*, Basil Blackwell.
- MENZIES, P. [forthcoming 2]: 'Chancy Causal Conditionals'.
- PETTIT, P. [1991]: 'Realism and Response-Dependence', *Mind*, 100, pp. 587–626.
- PRICE, H. [1986]: 'Against Causal Decision Theory', *Synthese*, 67, pp. 195–212.
- PRICE, H. [1988]: *Facts and the Function of Truth*, Basil Blackwell.
- PRICE, H. [1991a]: 'Agency and Probabilistic Causality', *British Journal for the Philosophy of Science*, 42, pp. 157–76.
- PRICE, H. [1991b]: 'Two Paths to Pragmatism', in P. Menzies (ed.), *Response-Dependent Concepts (Working Papers in Philosophy, No. 1)*, Philosophy Program, Research School of Social Sciences, ANU, pp. 46–82.
- RAMSEY, F. P. [1978]: *Foundations*, D. H. Mellor (ed.), Routledge & Kegan Paul.
- SMITH, P. [1987]: 'Subjectivity and Colour Vision', *Proceedings of the Aristotelian Society*, Supp. Vol., 61, pp. 245–64.
- SUPPES, P. [1970]: *A Probabilistic Theory of Causality*, North-Holland.
- TAYLOR, R. [1966]: *Action and Purpose*, Prentice Hall.
- WRIGHT, G. H. VON [1975]: *Causality and Determinism*, Columbia University Press.