

Transparent and Opaque Causation

© D. H. Mellor 1999

1 The opacity of intentional states

It is obvious that our beliefs about the world have causes and effects, both internal and external. Their most important external causes are the facts we perceive, i.e. the facts that our senses, when they are working properly, give us true beliefs about: these are the facts which, via our senses, cause us to believe in those very facts. Thus when, for example, I see Marion at the bar, what my eyesight does is enable the fact that Marion is there to cause me to believe that she is there, thus causing me to have that true belief about her. Similarly with the actions that are the external effects of our desires and beliefs, e.g. my wanting to buy Marion the drink I believe I owe her, a combination of desire and belief which causes me to offer her a drink. And so on. That many of our beliefs are caused in this way by our senses, and then combine with our desires to cause our actions, are among the platitudes of so-called 'folk psychology' that seem to me undeniable.

A common assumption about causation, however, makes these platitudes present a serious problem to which I want to offer a solution. The problem is caused by the so-called opacity of many beliefs, desires and other so-called 'intentional' states of mind. To illustrate this opacity, suppose that, in the 1998 German Federal election, some people voted SPD because they *wanted* Herr Schröder to become Chancellor, while others, given the SPD's economic policies, changed their investments because they *believed* he would become Chancellor. From this it does not follow that these people desired or believed that, as Chancellor, Schröder would run the country, since they might have desired or believed that his Finance Minister would run it for him.

Wanting or believing Schröder to be the Chancellor must therefore be, as it obviously is, a different state of mind from wanting or believing him to run the country, even if as Chancellor he does in fact run it. Moreover, the fact that these are different states of mind is not only obvious in itself, it also follows from their having different causes and effects. Thus while Marion's belief that Schröder is the Chancellor was no doubt originally caused by an announcement of the 1998 election result, her belief that he runs the country may have had to await the evidence of his government's actions. And while the first belief (but not the second) would affect how she would address any letter she sent to him, the second might well have more effect than the first on what she put in such a letter.

Why do these obvious facts about beliefs and desires show them to be opaque, and why should that opacity make their causation problematic? Taking these questions in turn, note first that if the Chancellor *does* run the country, the sentences 'Schröder is the Chancellor' and 'Schröder [is the man who] runs the country' cannot differ in truth value: that is, neither can be false if the other is true. In other words, if the singular terms 'the Chancellor' and 'the man who runs the country' refer to the same person, the true sentence 'Schröder is the Chancellor' cannot be made false by replacing the former term in it by the latter.

Any sentence with this property, of keeping its truth value (true or false) if a singular term ‘ x ’ in it is replaced by any other term that refers to x , is said to be *transparent for x* . ‘Schröder is the Chancellor’ has this property; but ‘Marion believes that Schröder is the Chancellor’ does not, since even if the Chancellor does run the country, Marion may not believe that he does. This is what is meant by calling that belief of hers opaque.

2 The transparency of causal relations

Why should the undeniable opacity of this and many other intentional states appear to conflict with the equally undeniable fact that they have causes and effects? To show why, I must say something about causation, by which I shall here mean not *general* causation (as in smoking causing cancer) but *singular* causation (as in Pat’s smoking causing Pat’s cancer). For obviously, if the desires and beliefs that each of us has at any given time *are* causes and effects, they are singular ones, as are the causes and effects they have. So if there is a problem with taking these states of mind to have causes and effects, it must lie in some assumed feature of singular causation. What is that feature?

The assumption that creates a problem for mental causation is neither *recherché* nor contentious. The problem does not for example arise from taking singular causes to be, in the circumstances, sufficient for their effects, or necessary for them, or any combination of the two (e.g. Mackie 1965). Nor, if we allow causation to be indeterministic, does it arise from assuming that causes must raise the chances of their effects (Mellor 1988). Nor does it depend on taking all causation to be transitive (Lewis 1973), or to instantiate laws of nature (Hempel 1965, §2.2), or on any of a host of other contentious assumptions about it (see Sosa and Tooley 1993).

What makes it hard to see how opaque mental states can be causes or effects is the idea that causation is, in some sense, a *relation* between causes and effects. And by ‘relation’ here I need not mean a *universal*, or a *trope*, or indeed anything that even nominalists who deny the existence of universals and tropes need deny (see Mellor and Oliver 1997, Introduction). For the problem I have in mind can be stated in terms not of relations but of relational *sentences*. These – sticking for simplicity to two-term relations – we may take to be sentences of the form ‘ aRb ’ (like ‘Pat smokes more than Fred’) which everyone, whatever they think relations are, can agree say that a relation R holds between two entities a and b . What are the essential features of such sentences?

Note first that we may take ‘ aRb ’ to be *atomic*, i.e. to contain no other sentences within it. Whether all relational sentences are in fact atomic depends on whether we let them include complex sentences like ‘ a runs Germany if b doesn’t’, which does contain other sentences (‘ a runs Germany’ and ‘ b doesn’t [run Germany]’). But even if we do admit such sentences, we can still reduce them to the form ‘ aRb ’ by abbreviating predicates like ‘... runs Germany if ... doesn’t’ to ‘ R ’: it will make no odds to the argument that follows.

The two assumptions that do matter to what follows are these.

- (i) *Existence*. For ‘ aRb ’ to be a relational sentence, it must entail that a and b exist, for the obvious reason that nothing which does not exist can be related to anything.

(ii) *Transparency*. A sentence ' aRb ' that is used simply to say that a relation R holds between a and b must be transparent for a and b . That is, its truth value cannot be changed by replacing either ' a ' or ' b ' with any other term referring to a or b .

Thus for 'Schröder runs Germany' to state a relation between Schröder and Germany, it must at least (i) entail the existence of, and (ii) be transparent for, both Schröder and Germany, as of course it does and is.

The transparency condition may be less obvious than the existence condition, but on reflection it is no less hard to deny. For consider any sentence of the form ' a is F ' which is just used to say that a has some property F . If F really is a property of a itself, and not of some part or aspect of a , its possession by a and hence the truth of ' a is F ' cannot possibly depend on how we refer to a . In other words, ' a is F ', so used, must be transparent for a . Similarly, if R really relates a and b , and not merely some parts or aspects of a and b , then the sentence ' aRb ' that says so must also be transparent for a and b . For if it was not, then whether R related a and b would depend on what we called them, which is absurd.

This is not of course to deny that a sentence of the form ' aRb ' can be opaque for a or b , or fail to entail their existence: of course it can. It is simply to deny that such sentences only (or even) ascribe relations to things. Take for example ' a believes that Lufthansa flight b lands at noon'. That does not state a relation between a and flight b , because it does not even entail that flight b exists, since it might have been cancelled. Nor does ' a believes Schröder is the Chancellor' state a relation between a and Schröder, even if it entails that Schröder exists: for while it is transparent for a , it is (as we have seen) opaque for Schröder, since even if Chancellor Schröder is the man who runs Germany, a may not believe that he is.

3 Causes and effects: events or facts?

What has all this to do with causation? Well, suppose we assume, as most philosophers do, that singular causation (if it exists) is a relation between singular causes and their singular effects. But then, for the reasons just given, it must be statable in sentences of the form ' aRb ', namely ' c causes e ', where c is the cause and e is the effect, which (i) entail that c and e exist and (ii) are transparent for c and e . Does ' c causes e ' always meet these two conditions for being a relational sentence?

The existence condition, (i), poses no problem for any singular causation, whether it links physical or mental causes and effects. Suppose for example I believe that the quickest way for me to reach Konstanz from England involves flying to Zurich. That belief of mine cannot have caused me to fly to Zurich unless (a) I had it and (b) I did fly to Zurich. And similarly in all other cases. Nothing that does not exist can cause or be caused by anything, and no one thinks it can.

The problem arises from (ii), the transparency condition. And to see how it does so, we must look more closely at what singular causes and effects are. There are two main candidates, *facts* and *events*, both of which I must now say a little more about, since the terms 'event' and 'fact' are used by different philosophers to mean very different things.

By *facts* I mean entities that correspond by definition (but not necessarily one-to-one) to

true, and only to true, sentences. In other words: if a sentence 'P' is true, it is a fact that P, i.e. the fact that P exists; if 'P' is not true, there is no such fact. Thus if Schröder does in fact run Germany, so that the sentence which says he does is true, then it is a fact that he does: that fact exists. If he does not, and 'Schröder runs Germany' is false, then that fact does not exist.

By singular or token *events* I follow Davidson (1970) in meaning a kind of *particular*, by which in turn I mean entities that (a) correspond to singular terms but not to sentences and (b) are ranged over by first-order rather than second-order quantifiers. (Condition (a) stops 'the fact that P' making facts count as particulars, while condition (b) stops properties of particulars, like *the colour red*, over which second-order quantifiers range, as in 'some colours are bright', also counting as particulars.) My flight to Zurich in December 1998 and Marion's voting in the 1998 election are obvious examples of events in this sense. And as all that matters here about events is that they are particulars, we may also take them to include any persistent *state*, like Marion's belief that Schröder is the Chancellor, when that is construed as a particular rather than a fact.

Although the existence of both events and facts, so understood, remains contentious, in what follows I shall take them both for granted, since I need only very weak assumptions about them. I need not for example assume that events are irreducible to less controversial particulars, such as people, as in the equation of 'My flight to Zurich occurred' with 'I flew to Zurich'. Similarly my facts need not be the substantial entities that make sentences like 'I flew to Zurich' true, merely the ontological shadows of true sentences generated trivially by the principle stated above, namely that 'P' is true if and only if it's a fact that P.

Supposing then, if only for the sake of argument, that there are events and facts in my weak senses, how do they differ? How for example does the fact that I flew to Zurich differ from my flight to Zurich? The answer is that the former is the *existential* fact that there is an event like the latter (Ramsey 1927 p. 37). What 'I flew to Zurich' says is that there is (or was in December 1998) a particular *event* of a certain sort, namely a flight by me to Zurich; just as 'Germany has a Chancellor' says that there is (now) a *person* of a certain sort, namely a Chancellor of Germany. And just as the particular person, Schröder, who is Germany's Chancellor must not be confused with the fact that Germany has a Chancellor, so the particular event, my flight to Zurich, must not be confused with the fact that I flew there.

Given this distinction, are causes and effects events or facts? If they are events, then causation will be most perspicuously reported by relational sentences of the form '*c* causes *e*', where 'causes' is a predicate and '*c*' and '*e*' are singular terms referring to those particular events. While if causes and effects are facts, the most perspicuous way to report causation will be by using the connective 'because' to link true sentences stating the cause and its effect: as in 'I flew to Zurich because I believed that was the quickest way to reach Konstanz'.

Thus used, 'because' differs from other well-known connectives, like 'and' and 'or', by not defining a complete truth function (Davidson 1967, §I). That is, the truth value (*true* or *false*) of a causal 'E because C', unlike that of 'C and E', is not always fixed by the truth

values of 'C' and 'E'. In particular, it is not fixed when 'C' and 'E' are true: for if it was, then 'E because C' would have to be either always false or always true for *all* true 'C' and 'E'. But if it was always *false*, there would be *no* factual causation, which by hypothesis there is; while if it was always true, then all facts would cause all other facts, which of course they don't. So for example 'E because C' could only be a complete truth function if either 'I flew to Zurich because I believed that was the quickest way to reach Konstanz' was false, which it isn't, or its converse, 'I believed flying to Zurich was the quickest way to reach Konstanz, because I flew to Zurich', was true too, which it isn't.

However, even if the molecular sentence 'E because C' is not a *complete* truth function of the sentences 'C' and 'E' it contains, it is at least a *partial* truth function of them. That is, some truth values of 'C' and 'E' do fix its truth value. For if either 'C' or 'E' is false, so is 'E because C'. If I didn't fly to Zurich at all, I couldn't have flown there *because* of anything; nor could I have done anything *because* I believed that flying to Zurich was the quickest way to reach Konstanz if I didn't in fact believe that. In other words, for all 'C' and 'E', 'E because C' entails both 'C' and 'E'. And this, on my weak definition of facts, means that only if C and E are facts can 'E because C' be true.

4 Opaque mental causation

Does the entailment of 'C' and 'E' by 'E because C' make factual causation relate facts in my weak sense of 'fact'? We can certainly represent it as doing so, by restating 'E because C' as 'The fact that C causes the fact that E' or, for brevity, 'C causes E', thus turning it into the canonical form of '*c* causes *e*'. And this restatement meet at least one of the two conditions on relational sentences given in **2**, namely that 'C causes E' must entail the existence of the facts C and E. For in my sense of 'fact', all this means is that the sentence 'E because C' entails that the sentences 'C' and 'E' are true, which everyone will agree that it does.

The real question is whether 'C causes E', i.e. 'E because C', meets the *transparency* condition on relational sentences, namely that no replacement of a singular term within it by another term referring to the same entity will change its truth value. This is the condition which Davidson (1967, §I) has argued cannot be met unless 'E because C' is a *complete* truth function of 'C' and 'E', which we have just seen that it cannot be. That argument is both complex and controversial, and I think it is unsound (Mellor 1995, ch. 9.4–5). But for present purposes I need not argue the point: for we don't need it to show that 'E because C' cannot be transparent when C or E is a fact about what someone believes or desires, as our earlier example will serve to show.

Suppose Chancellor Schröder is in fact the man who runs Germany. Even so, as we have seen, Marion's belief that he is the Chancellor differs from her belief that he runs Germany, and may have different causes and effects, such as those mentioned in **1**. Suppose it does. Then while 'Marion believes Schröder is the Chancellor, because she saw the election results' is true, 'Marion believes Schröder [is the man who] runs Germany, because she saw the election results' is false. So there are opaque truths of the form

'Marion believes Schröder is the Chancellor, because C',

just as there are of the form

‘E, because Marion believes Schröder is the Chancellor’

Moreover, Marion could believe that Schröder is the Chancellor but *not* that he runs Germany, even though he does. So even if she has both beliefs, and even if they share some causes and effects, sentences saying what those causes and effects are would still be made opaque by the opacity of ‘Marion believes Schröder is the Chancellor’ and the fact that ‘E because C’ entails both ‘C’ and ‘E’.

Hence our problem. Because the opacity of beliefs, desires and other intentional states infects sentences stating their causes and effects, those sentences cannot have the transparency that sentences stating any relation, causal or otherwise, must have. But then it seems as if facts about my beliefs and desires can neither cause nor be caused by any other facts. I cannot really have flown to Zurich *because* I had certain beliefs or desires; I cannot really believe that Schröder is Chancellor, *because* someone told me he was; and so on. Yet in all these cases clearly there are *some* mental causes and effects; so if they cannot be facts about what I believe and desire, they must be particular events or states, i.e. entities for which, precisely because they do *not* correspond to sentences with truth values, questions of transparency and opacity do not arise.

But this conclusion is most unsatisfactory in at least two respects. First, instead of letting my beliefs and desires be properties of a single person – me – it turns them into different particulars – events or states – whose relations to each other and to me are then left quite unexplained. Second, it allows the mental, as opposed to the physical, properties of such an event or state – e.g. the fact that it is a belief or a desire, with such-and-such a content – to be irrelevant to its causes and effects, a consequence to which philosophers have often and rightly objected (Crane 1992). For it is surely not enough to say that what caused me to fly to Zurich was a state which *is* my belief that this was the quickest way to reach Konstanz, if what made that state cause me to do this was *not* this fact about it but the fact that it had some quite different non-intentional physical property.

5 Transparent factual causation

We must therefore look more closely to see if the transparency of relational sentences really does stop facts in general, and especially facts about what we believe and desire, being causally related to other facts. And as we shall see, appearances to the contrary, it doesn't: causation can indeed be a relation between facts, including facts about our intentional states. But to show how, I must first make clearer what the transparency of causation between facts does, and does not, require.

All transparency really requires here is that statements of the form ‘The fact that C causes the fact that E’, which I have abbreviated to ‘C causes E’, be transparent for the facts C and E. That is, replacing the singular term ‘C’ or ‘E’ with any other term referring to the same fact must not make a true ‘C causes E’ false or a false one true. But then, having taken ‘C causes E’ to be equivalent by definition to ‘E because C’, in which ‘C’ and ‘E’ are not singular terms but sentences, the transparency requirement becomes extended from terms that refer to the

facts C and E to terms within the sentences ‘C’ and ‘E’ that refer to *particulars* (Davidson 1967, §I).

But why should the transparency requirement be extended in this way? Why should we assume that replacing one co-referring term by another within a true sentence never changes the fact to which that sentence corresponds? For it isn’t obvious that it doesn’t: on the contrary, in many cases it is obvious that it does. Thus suppose again that Schröder runs Germany, so that the terms ‘Schröder’ and ‘the man who runs Germany’ refer to the same person. Does this really entail that Schröder’s being the Chancellor is the same fact as the fact that he runs Germany?

Of course not. These facts – which are not facts about intentional states – are obviously quite different, with quite different causes and effects. But then the sentence ‘C causes E’ can quite well be transparent for the facts C and E even if – indeed precisely because – the sentence ‘E because C’ is opaque for particulars referred to within the sentences ‘C’ and ‘E’. For if ‘Schröder is the Chancellor’ and ‘Schröder runs Germany’ state different facts, those facts can easily have different causes and effects, statable in sentences that can be transparent for those facts even if they are opaque for the Chancellor, i.e. for Schröder.

Why then should we resist the obvious conclusion that Schröder’s being the Chancellor and Schröder’s running Germany are two different facts? The usual reason is that distinguishing these two facts makes the concept of a fact *non-extensional*. That is, it stops the identity statement ‘ $a=b$ ’ and the obvious transparency of ‘ a is F ’ entailing that a ’s being F is the same fact as b ’s being F . But why should we accept this entailment in the face of such clear counter-examples as the one I have just given (where ‘ a ’ is ‘the Chancellor’, ‘ b ’ is ‘the man who runs Germany’ and ‘ F ’ is ‘Schröder’)? The reason I think is a desire to preserve the simple criterion of identity for facts which the entailment gives us.

6 Identity criteria for causes and effects

How strong is this objection to a non-extensional concept of a fact? We can after all accept the Frege–Quine slogan ‘no entity without identity’ without accepting ‘no entity without identity *criteria*’. If criteria of identity are needed at all, they are needed, not by the entities whose identity criteria they are, but by us, to enable us to talk about those entities. But do we really need criteria of identity for entities in general, and for facts in particular, in order to talk about them? I say not; and we certainly do not need a *single* criterion of identity for facts of all kinds. After all, we have no such criterion for *particulars* of all the many uncontroversial kinds we talk about, ranging from galaxies through planets, nations, people, plants and cells, down to atoms and quarks. And if no one demands a single criterion for all these diverse particulars, why should we demand one for all facts, however diverse – or, come to that, for all events?

Yet many philosophers have demanded such a criterion for all causally related events. And despite thinking the demand unreasonable, I propose to avoid controversy by accepting it, and applying to all causally related *facts* the very same causal criterion of identity that Davidson (1969) originally offered for his causally related *events*.

The causal criterion of identity for events may be stated as follows. An event d that has any causes or effects is identical to an event $d^{\mathbb{W}}$ if and only if d and $d^{\mathbb{W}}$ have all the same causes and effects. That is, replacing the singular term ' d ' by the singular term ' $d^{\mathbb{W}}$ ' never makes a true ' c causes d ' or ' d causes e ' false or a false one true. Similarly, I say, for facts (Mellor 1995, ch. 9.3). A fact D that has any causes or effects is identical to a fact $D^{\mathbb{W}}$ if and only if D and $D^{\mathbb{W}}$ have all the same causes and effects. That is, replacing the singular term ' D ' by the singular term ' $D^{\mathbb{W}}$ ' never makes any true ' C causes D ' or ' D causes E ' false or a false one true. Or, put in terms of sentences, replacing the sentence ' D ' by the sentence ' $D^{\mathbb{W}}$ ' never makes any true ' E because D ' or ' D because C ' false or a false one true.

This identity criterion is by no means uncontentious, either for events or for facts. But it is at least as good for facts as it is for events. Indeed it is better: for with it, we can use transparent relational sentences to state the causes and effects of all sorts of facts, including facts that we need opaque sentences to state, like the fact that Schröder is the Chancellor.

7 Mental causes and effects

And as for the fact that Schröder is the Chancellor, so for the fact that Marion believes he is. Just as, by our causal criterion, the former fact must differ from the fact that Schröder runs Germany, because that fact has different causes and effects, so the latter fact must differ from the fact that Marion believes he runs Germany, because that fact too has different causes and effects. So these two pairs of facts are on a par: we can use transparent relational sentences to state the causes and effects of each member of both pairs in exactly the same way. Yet the second is a pair of intentional facts, about what Marion believes, and the first is not. So whatever other problems may be posed for the philosophy of mind by the obvious fact that our beliefs, desires and other intentional states have causes and effects, none is posed by their opacity; since this is matched by many statements of the causes and effects of non-mental facts, like the fact that Schröder is the Chancellor. If, therefore, as we have seen, those non-mental facts can have transparent causal relations, so too can facts about our beliefs and desires.

In short, by giving a uniform account of both intentional and non-intentional causes and effects, the theory that causation relates facts rather than particulars solves a long-standing and otherwise intractable problem of how intentional states can have causes and effects. And by so doing, it destroys a major argument for physicalism, namely that only non-intentional physical states of people can really have causes and effects (see Crane and Mellor 1990, §3). As that is false, our apparently intentional mental states need neither lack their apparent causes and effects nor be identified with non-intentional physical states. For that, if causation relates facts, is a false dichotomy: the opacity of our beliefs and desires is no obstacle to their having, in their own right, all the causes and effects they seem to have.

References

Crane, T. (1992), 'Mental Causation and Mental Reality', *Proceedings of the Aristotelian*

- Society* **92**, 185–202.
- Crane, T. and Mellor, D. H. (1990), 'There is No Question of Physicalism', *Mind* **99**, 185–206.
- Davidson, D. (1967), 'Causal Relations', *Journal of Philosophy* **64**, 691–703.
- Davidson, D. (1969), 'The Individuation of Events', *Essays in Honor of Carl G. Hempel*, ed. N. Rescher, Dordrecht: Reidel, 216–34.
- Davidson, D. (1970), 'Events as Particulars', *Noûs* **4**, 25–32.
- Hempel, C. G. (1965), 'Aspects of Scientific Explanation', *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: The Free Press, 331–496.
- Lewis, D. K. (1973), 'Causation', *Journal of Philosophy* **70**, 556–67.
- Mackie, J. L. (1965), 'Causes and Conditions', *Causation*, ed. E. Sosa and M. Tooley, (1993), Oxford: Oxford University Press, 33–55.
- Mellor, D. H. (1988), 'On Raising the Chances of Effects', *Probability and Causality*, ed. J. H. Fetzer, Dordrecht: Kluwer, 229–39.
- Mellor, D. H. (1995) *The Facts of Causation*, London: Routledge.
- Mellor, D. H. and Oliver, A., eds (1997), *Properties*, Oxford: Oxford University Press.
- Ramsey, F. P. (1927), 'Facts and Propositions', *Philosophical Papers*, (1990), Cambridge: Cambridge University Press, 34–51.
- Sosa, E. and Tooley, M., eds (1993), *Causation*, Oxford: Oxford University Press.