



OXFORD JOURNALS  
OXFORD UNIVERSITY PRESS

## Mind Association

---

Can Edgington Gibbard Counterfactuals?

Author(s): Adam Morton

Reviewed work(s):

Source: *Mind*, New Series, Vol. 106, No. 421 (Jan., 1997), pp. 101-105

Published by: [Oxford University Press](#) on behalf of the [Mind Association](#)

Stable URL: <http://www.jstor.org/stable/2254773>

Accessed: 11/10/2012 15:42

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at  
<http://www.jstor.org/page/info/about/policies/terms.jsp>

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.



Oxford University Press and Mind Association are collaborating with JSTOR to digitize, preserve and extend access to *Mind*.

<http://www.jstor.org>

## *Can Edgington Gibbard Counterfactuals?*

ADAM MORTON

Dorothy Edgington's interesting and valuable survey article on the current state of work on conditionals (Edgington 1995) contains several arguments for the conclusion that few non-material conditionals have truth values. The case is much stronger for "indicative" than for so called subjunctive or counterfactual conditionals. One argument of Edgington's would be particularly telling if it succeeded. She argues that the Gibbard phenomenon, which provides a hard to resist case for denying truth values to many indicative conditionals, can be extended to counterfactuals. The purpose of this note is to express a doubt about the argument, one which Edgington should address if she is to convince us.

Edgington's example of the Gibbard phenomenon (see Gibbard 1981, also Jackson 1990, 1991 and Lowe 1991) is as follows. Call it the live-Jones case. There is a disease *D*, vaccines *A* and *B*, and a side-effect *S*. Neither *A* nor *B* alone completely prevents *D*. If you've had *A* and you go on to get *D* you get *S*; but if you've had *B* and you go on to get *D* you don't get *S*. If you've had both *A* and *B* you don't get *D* and so don't get *S*. There are two observers *X* and *Y* and a patient, Jones. *X* knows that Jones has had *A* and thus is justified in believing that *if Jones gets D he will get S*. *Y* knows that Jones has had *B* and thus is justified in believing that *if Jones gets D he will not get S*. Each of their beliefs is justified by what they know. They contradict one another, but learning the whole truth will not show that one is right, since the whole truth includes the fact that Jones has had both *A* and *B* and thus will not get *D*. The conclusion Edgington and others draw is that the two conditionals cannot have truth values.

I take the reason for thinking that they cannot have truth values to run along the following lines. The facts are symmetrical between them, in that there are equally good reasons for thinking that one is true as that the other is. So one is true iff the other is. Call this *Symmetry*. But they contradict one another: if getting *S* is a consequence for Jones of getting *D* then escaping *S* is not a consequence. Call this *Contradiction*. So the one is true iff the other is not. But these two biconditionals are contradictory. (Note that they can be contradictory even if the sentences they discuss have no truth value.) So we had better not give any truth values. This reasoning is not completely uncontroversial (see Jackson on this). But I shall not challenge it.

That is the Gibbard phenomenon. Edgington goes on to consider a dead-Jones case. Suppose Jones is run over by a bus before there is any chance of his getting *D*. Then, she argues, *X* can say *if Jones had got D he would have got S* and *Y* can say *if Jones had got D he would not have got S*. As a result “at that time, the Gibbard phenomenon applies—each has adequate reason for his opinion, and the world rules out there being an objectively correct opinion, for it rules out Jones’ getting the disease” (Edgington 1995, p. 319).

It is not completely clear what the claim is here. It cannot be that the counterfactuals have no objective truth value just because the world rules out Jones’ getting the disease, for there are plenty of counterfactuals with impossible antecedents (“if we were to flap our arms and fly to the top of that building, we could see the sun set”). This parallels the fact that the Gibbard phenomenon is not that when we know that the antecedent of an indicative conditional is false we know there is no correct opinion as to its truth. For in the standard example we *know* that Oswald shot Kennedy, and also know that if he did not someone else did. I shall take the claim to be that when one has a Gibbard example for a pair of indicative conditionals, as in the live-Jones case, a slight modification can transform it into a Gibbard example for a corresponding pair of counterfactual conditionals, as is claimed of the dead-Jones case, without tampering with the facts that make the indicative version a Gibbard case. *Symmetry* and *Contradiction* can be produced for the counterfactuals without specifying anything that makes them cease to hold for the indicatives. Thus Edgington has to make two claims about the dead Jones case. First, that the case can be spelled out so that there are no further facts which favour one counterfactual over its contrary which do not also favour one of the indicatives in the live Jones case over its contrary. (I shall use “contrary” as a neutral term, hoping not to beg any questions about truth values and contradictions.) And second that in the situation thus spelled out either of the two counterfactuals is true iff the other is false.

It is easy to think of ways in which the two counterfactuals in the dead-Jones case could apparently have opposite truth values. Suppose that Jones almost missed the appointment to have vaccine *A* but vaccine *B* was administered at a time and place at which only really extraordinary circumstances would have prevented Jones’ presence. Then if Jones had got *D* it would have been because he did not get *A*, so that “if he had had *D* he would not have got *S*” is true, and “if he had had *D* he would have got *S*” is false.

But does this story preserve the facts that make the corresponding live-Jones case a Gibbard case? The only difference is that Jones is not run over by the bus. The facts about the inoculations can stay the same, and

still one observer can know that he did have *A* and another that he did have *B*. Observer *X* has grounds for saying “if he had not been run over I would have asserted ‘if he gets *D* he will get *S*’”. So when they put their information together and learn that he did have both they will conclude that both indicatives would have been false. They can thus conclude “if he had not been run over we would now be denying both ‘if he gets *D* he will get *S*’ and ‘if he gets *D* he will not get *S*’”.

So there is one set of facts which preserves Gibbard for indicatives and not for counterfactuals. Might there not be another way of spelling out the facts which is friendlier to Edgington’s claim? Suppose that Jones gets inoculations *A* and *B* before birth. Suppose in fact that they are carried to him at conception by his parents’ sperm and egg, so that an uninoculated Jones can hardly be taken to be Jones. Then *Symmetry* is assured, as is the falsity of the antecedent of the counterfactuals. *Contradiction* is now crucial. Are “if Jones had got *D* he would have got *S*” and “if Jones had got *D* he would not have got *S*” contradictories? If not, it is as plausible to say that they are both false as that they both lack a truth value.

One is certainly not forced to deny truth values here. In the indicative case one can reason as follows. Suppose that Jones gets *D*. Then either he actually will get *S* or he actually will not. So “if *D* then *S*” and “if *D* then not *S*” are contradictories, and not just contraries. But in the counterfactual case this doesn’t work. Suppose Jones had got *D*. Then whether or not he actually got *S*, whether he would have got *S* depends on what happens in the counterfactual situations in which he got *D*. Typically in some he will get *S* and in some he will not. (Perhaps in some he gets *D* because a mutant variety of the disease is unaffected by *B*, and thus gets *S*; perhaps in some he gets *D* because a mutant variety is unaffected by *A*, and thus does not get *S*; perhaps...) If neither the *S* ones nor the not-*S* ones are more likely (or closer, or more accessible) then both “If Jones had got *D* he would have got *S*”, and “If he had he wouldn’t” can be false. The important point is not that the standard analyses of counterfactuals block conditional excluded middle, but that the reasoning that makes it plausible for indicatives fails for counterfactuals. So whether or not one accepts those standard analyses, one is not forced to take counterfactuals with contradictory consequences as contradictories.

(For a discussion of why conditional excluded middle—if *p* then *q* or if *p* then not *q*—does not hold for counterfactuals see Lewis 1973, pp. 79–82. My own view is that conditional excluded middle, for counterfactuals, has what plausibility it does only because English is so hopeless about the scopes of negations.)

Perhaps we have not made the antecedent impossible enough. Since in a Gibbard case information about the antecedent can accumulate it is not

plausible to make the antecedent absolutely impossible, but it can be impossible given surrounding circumstances. In an earlier article Edgington describes a case in which the antecedent is “the temperature was  $T$ ” in a situation in which the gas laws prohibit it from being  $T$  given what we know about the pressure and volume (see Edgington 1992, pp. 206–7). Then we might suppose that the world rules out the temperature being  $T$  so firmly that we cannot assign any value to counterfactuals beginning “if the temperature had been  $T$ ”. But exactly the same replies are possible as before. We can consider all the ways in which the pressure or volume could have been different, thus allowing the temperature to be  $T$ , even if these involve unlikely chains of events beginning far back in the past. If these favour a consequent over its negation then one counterfactual is true and its contrary false. And if there is a symmetry between the two, then again it is at least as plausible to say that both are false as that both lack truth value.

My conclusion is that one can respond to Gibbard-like situations involving counterfactuals by judging the contrary counterfactuals to be both false instead of neither true nor false. Must one respond in this way? One reason for avoiding truth-valuelessness with counterfactuals is the greater range of facts that are potentially relevant. A body of evidence often definitely entails or refutes an indicative conditional, so that for example if one knows the medical facts and knows that Jones has had inoculation  $A$ , one knows that if he gets  $D$  he will get  $S$ . This is rarely the case with counterfactuals. A body of evidence can support a counterfactual but rarely entails or refutes it. And an expanded evidence set can support a contrary counterfactual. We learn that there is a mutant version of  $D$  that only affects people inoculated at conception; we learn that a technician narrowly avoided installing a faulty pressure-monitoring device. Then the evidence still entails the same indicatives, but supports different counterfactuals. We are rash therefore ever to completely write off any counterfactual; we want to give most of them a credence somewhere between confidence of truth and confidence of falsity. But to deny them truth values is to take them off the true/false scale completely.

There is one more reply to consider. Suppose there are facts that make one counterfactual true and the other false. Might knowledge of these facts not be used to break the symmetry that makes the corresponding indicative case? (So then in proving that the counterfactuals were not Gibbarded one would have simply shown that the indicatives wouldn't Gibbard in that case.) Might we not say that if we know that Jones could quite easily have missed his  $A$  inoculation then we know that if he gets  $D$  he will not get  $S$ ? I find this contrary to my sense of the indicative conditional, since we know that in fact Jones will not get  $D$ . But allow it. Then we have

one fewer Gibbard case for indicatives. Since Gibbard cases for counterfactuals are hard to come by, this strategy threatens to dry up the supply of them for indicatives.

I conclude that Gibbard symmetries between facts about what is do not extend in the required way to symmetries about what could have been. Symmetries in fact are consistent with asymmetries in possibility. And asymmetries in possibility are much more likely to result in falsity of both contrasting counterfactuals. The underlying question here is whether the plausibility of conditional excluded middle for indicatives extends to counterfactuals; it is on that point that Edgington owes us some more explanation.

*Department of Philosophy*  
*University of Bristol*  
*9 Woodland Road*  
*Bristol BS8 1TB*  
*UK*  
*adam.morton@bristol.ac.uk*

ADAM MORTON

#### REFERENCES

- Edgington, Dorothy 1992: "Matter-of-fact conditionals". *Proceedings of the Aristotelian Society*, supp. vol., 65, pp. 185–210.
- 1995: "On conditionals". *Mind*, 104, pp. 235–330.
- Gibbard, Allan 1981: "Two recent theories of conditionals", in W.L. Harper et al. (ed.), *Ifs*. Dordrecht: Reidel, pp. 211–47.
- Jackson, Frank 1990: "Classifying conditionals". *Analysis*, 50, pp. 134–47.
- 1991: "Classifying conditionals II". *Analysis*, 51, pp. 137–43.
- Lewis, David 1973: *Counterfactuals*. Oxford: Blackwell.
- Lowe, E. J. 1991: "Jackson on classifying conditionals". *Analysis*, 51, pp. 126–36.