Meeting of the Aristotelian Society held at Senate House, University of London, on 8 October 2007 at 4:15 pm.

The Presidential Address

I— Counterfactuals

Dorothy Edgington

I argue that the suppositional view of conditionals, which is quite popular for indicative conditionals, extends also to subjunctive or counterfactual conditionals. According to this view, conditional judgements should not be construed as factual, categorical judgements, but as judgements about the consequent under the supposition of the antecedent. The strongest evidence for the view comes from focusing on the fact that conditional judgements are often uncertain; and conditional uncertainty, which is a well-understood notion, does not function like uncertainty about matters of fact. I argue that the evidence for this view is as strong for subjunctives as it is for indicatives.

I

Introduction. 'The analysis of counterfactual conditionals', wrote Nelson Goodman in 1946, 'is no fussy little grammatical exercise. Indeed, if we lack the means for interpreting counterfactual conditionals, we can hardly claim to have any adequate philosophy of science' (Goodman 1965, p. 3). Since then, they have figured in philosophy more widely than in the philosophy of science: in accounts of causation, perception, knowledge, rational decision, action, explanation, and so on. And outside philosophy, in ordinary life, counterfactual judgements play many important roles, for instance in inferences to factual conclusions: 'It's not a problem with the liver', says the doctor; 'for the blood test was normal; and if it had been a problem with the liver it would have been [such-and-such].' Yet, as I shall try to show, the most widely accepted theories of counterfactual conditionals face serious problems; and I shall advocate a different approach.

There is philosophical dispute about whether a single theory of conditionals can handle both indicative and subjunctive condition-

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

als,[1] or whether the differences between them are such that the theory of conditionals splits into two at the outset. What is the relationship between 'If she catches the four o'clock train she will be here by five' and 'If she had caught the four o'clock train she would have been here by five'? Those who defend the textbook truth-functional theory of the former, indicative conditional must see the latter, 'subjunctive' or 'counterfactual' conditional as different, and indeed logically stronger than the indicative conditional—as Frank Jackson (1987) and David Lewis (1973, 1976) do. Those who take the ordinary indicative conditional also to be logically stronger than material implication (the truth-functional conditional), cite in their favour the advantage of being able to provide a broadly unified theory of conditionals, within which one hopes to explain the difference made by 'the combination of tense, aspect and mood that we have gotten into the habit of calling "subjunctive"' (these words are from Robert Stalnaker (2005), and he advocates a unified approach). This perspective goes back a long way. It informs P. F. Strawson's writings both in *Introduction to Logical Theory* (1952) and in a later paper on H. P. Grice: Strawson considers pairs like those above, and says 'The least attractive thing to say about the difference between [the members of such a pair] is that "if" has a different meaning in one from its meaning in the other' (1986, p. 230).

I want to look at this issue from the perspective of a suppositional account of conditionals. On this view, a conditional statement is not a categorical assertion of a proposition, true or false as the case may be; it is rather a statement of the consequent under the supposition of the antecedent. A conditional belief is not a categorical belief that something is the case; it is belief in the consequent in the context of a supposition of the antecedent. The strongest evidence for this view comes from considering uncertain conditional judgements. No one can deny that our conditional judgements are often uncertain; and our best account of this—of conditional uncertainty, uncertainty about the consequent under the supposition of the antecedent—does not behave like uncertainty about the truth of a proposition. My topic is whether this perspective extends to counterfactuals.

As a thesis about indicative conditionals the suppositional theory is now quite popular, and accepted, or at least taken as a serious con-

[1] I use 'subjunctive' and 'counterfactual' interchangeably. As is widely recognized, neither term is ideal.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

tender whose virtues need to be accounted for, by many philosophers. But its extension to counterfactuals is very unpopular. Although Ernest Adams, who did most to make this theory of indicative conditionals fly, did extend it to subjunctives (Adams 1975, ch. 4; 1993), this part of his work has received little attention. Most of those who have followed Adams's lead on indicatives do not think of subjunctives in these terms—for instance, Anthony Appiah (1985), Allan Gibbard (1981), and Jonathan Bennett in his influential book (2003). When Bennett turns to subjunctives, an entirely new battery of machinery arrives on the scene, and he says 'they are not in any deep way like indicatives' (p. 256). Here he echoes Gibbard, who wrote that 'the apparent similarity of these two "if" constructions hides a profound semantic difference' (1981, p. 211).

Some opponents of the suppositional view argue that indicatives and subjunctives are closely linked, the suppositional view is obviously wrong for subjunctives, hence it must be wrong for indicatives too, for instance Jonathan Lowe (1995), adding in a footnote that Edgington does not accept this obvious fact about subjunctives, but 'few would go to that extreme'; Stalnaker, more tentatively, in conversation; and William Lycan, in his book *Real Conditionals* (2001) and in his review of Bennett's book (2005), makes this point. Having said that the propositional, truth-conditional account of subjunctives is 'contested by Edgington but otherwise uncontroversial', Lycan points out that according to Bennett, 'indicatives not only differ in meaning from their corresponding subjunctives, they do not even have the same *kind* of meaning' (2005, p. 118). (The latter express propositions, true or false as the case may be; the former do not express propositions. This is, indeed, a big difference.)

I am on Lycan's side in this matter: I think the indicatives and subjunctives are too closely linked for the Gibbard-Bennett view to be plausible. But I jump the other way. I try to make the case here that when we consider the uncertain judgements we express as subjunctive conditionals, the case is just as strong as it is for indicatives, that these conditionals do not express propositions—that they do not have truth conditions. And I try to do this better here than I have managed to do before.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

II

Judgements Involving Uncertainty. Amongst our judgements involving uncertainty, there is a kind which is helpful for understanding counterfactuals: a judgement to the effect that it *was* probable that such-and-such *would* happen. (But, one might consistently add, it didn't happen, or it is now unlikely that it did.) What does such a judgement mean? It doesn't mean that you were close to certain that such-and-such would happen. You might or might not have been. You can say 'It was probable that such-and-such would happen, though I didn't realize it at the time.' Indeed, you may be talking about a time before you were born, or even a time before anyone was born—about, say, the survival chances of dinosaurs. You can be expressing the opinion that there was, at the time in question, a high objective chance that such-and-such would happen. In any case, you are endorsing the corresponding hypothetical degree of confidence as (in your view) the right one at the earlier time.

Conversely, when I realize that my earlier high confidence was misplaced—for example, when I thought I would make money by investing in the swindler's project—I don't felicitously say 'It was very probable that I would make a fortune', but only 'I thought it was very probable that I would make a fortune', for I now do not endorse the earlier judgement.

These sorts of judgements, amongst others, can occur within the context of a supposition. We make suppositions when deliberating about what to do, and we make suppositions when deliberating about what is the case. Suppose it rains tomorrow; then we'll stay at home and watch TV. Suppose Smith took the money; then Mary was lying when she said he was with her. Suppose they had been at home; then the lights would have been on; but they aren't. Suppose the prisoner had jumped from that window; then the flowers below would have been squashed; and so they are. Suppose Kennedy had not been assassinated; then the Vietnam war would not have escalated so. You can be more or less certain of these judgements in the context of a supposition: that Jane will accept on the supposition that she is offered the job; that Smith took the money on the supposition that Jones didn't; that the dog would have bitten me on the supposition that I had approached, and so on. And suppositions, it is natural to think, are more succinctly formulated with the word 'if'. Note that the easy transitition between 'suppose' and 'if' is as

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

evident for subjunctives as it is for indicatives.

On the suppositional account, our backward-looking and forward-looking indicative conditional judgements express our degree of confidence in the consequent on the supposition of the antecedent. With the backward-looking ones, we all have our own idiosyncratic combinations of knowledge and ignorance, and, as Gibbard (1981) made famous, people can faultlessly come to opposite opinions: there is nothing objective to aim at. For instance, we both start off knowing that *X* or *Y* or *Z* did it. I discover it wasn't *Y*. You discover it wasn't *Z*. I accept 'If not *X*, *Z*' and reject 'If not *X*, *Y*'. You do just the opposite. All the relevant facts are available, but neither of us has them all, and we would know if we pooled our information that *X* did it. The present objective chance that *X* didn't do it is zero, so there is no such thing as the present objective chance of *Y* supposing that not *X* (just as there is no such thing as an objective chance that you will pick a spotted ball if you pick a red ball, when there are no red balls in the bag). For many forward-looking indicatives, by contrast, it is not yet determined whether the antecedent is true (or at least we treat it as not yet determined, not yet knowable, whether the antecedent is true), and there may be such a thing as the objective conditional chance of *C* given *A* for our judgements to aim at. And after the event, should it turn out that not *A*, we can be right or wrong when we say that, very likely, *C* would have happened if *A* had—very likely, you would have picked a spotted ball if you had picked a red ball (assuming now that there were some red balls in the bag). Thus, I hope, we get a fundamentally unified account of conditional judgements, which also explains interesting differences between the different kinds. Of course we may be certain that if *A*, *C*, but typically we are not, and conditional probability—the probability of *C* on the sup-position that *A*—is, I claim, the key to how close to certain we are of a conditional of any kind. For subjunctive conditionals, this conditional probability does not (normally) represent your current degree of belief in *C* given *A*, but (most typically) your view about how likely it *was* that *C would have* happened, given that *A* had.[2]

Pairs like the following have been supposed to force a separate

[2] The provisos are there because (a) sometimes, about the future, it might be a matter of indifference whether you say 'If it rains…' or 'If it were to rain…'; and (b) not every context shift away from your actual epistemic state is a temporal one: 'Suppose Euclidean geometry had been true of the actual world …'

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

account of subjunctives, by highlighting their difference from indicatives:

> If Oswald didn't kill Kennedy, someone else did;
> If Oswald hadn't killed Kennedy, someone else would have.

One may accept the former and reject the latter. But this is easily explained on the present view. It is obviously consistent to be very confident that someone else did it on the supposition that Oswald didn't, but to judge that it was very unlikely, back in 1963, that someone else would have killed Kennedy, supposing that Oswald hadn't.[3]

Why do these judgements matter, judgements about what would have happened if …? Here is one reason which I am inclined to think is the principal one: they play an indispensable role in empirical reasoning about what is the case. In abduction, or inference to the best explanation, we look for hypotheses such that what we do observe is what we *would* expect to observe if it were the case that *H* and would not expect to observe were it not the case that *H*. 'It's not a problem with the liver', says the doctor; 'for the blood test was normal; and if it had been a problem with the liver, it would have been [such-and-such]'; 'They're not at home; for the lights are off; and if they had been at home, the lights would have been on'; 'I think the patient took arsenic; for he has such-and-such symptoms; and these are the symptoms he would have if he had taken arsenic'; 'I think the prisoner jumped from that window; for the flowers below are squashed; and they would have been squashed if he had jumped from there.' These are not intended as deductively valid arguments. They can be defeated if reasonable alternative hypotheses make it just as likely, or unlikely, that we would have observed what we do observe. For instance it may be pointed out that the flowers would also be squashed if there had been a game of football, or a dog fight. Or it could be pointed out that they always leave the lights on when they go out at night, so there must be some other explanation of the lights being off—they have gone to bed early, or there was a power cut. Nevertheless, they are part and parcel of the most basic kind of empirical reasoning. Nor are the conditionals in-

[3] It is worth pointing out that, by the same token, one can be highly confident that if Oswald hadn't, no one else would have, but not at all confident that if Oswald didn't, no one else did. Theories according to which the subjunctive conditional is stronger than the indicative conditional can explain one divergence, but cannot give a symmetric explanation of the other.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

volved typically certain rather than probable. But to the extent that we can find a hypothesis H such that the chance was high that we would observe E (as we actually do), on the supposition that H is true, and the chance was low that E given $\neg H$, and H is not initially too unlikely, we have a good argument for H.

Here is another kind of inference of some importance. We all ate the same pie for dinner. Fred got ill and died. The question is whether the cause of death was food poisoning from the pie. No, it is argued, if Fred had been food-poisoned, we would have suffered the same fate (or at least this is quite likely) and we didn't.[4] We test a drug on rats, and conclude that it is safe for humans, for if it had such-and-such effects on humans, it would also have had these effects on the rats. We conclude that the projected dam is of a safe design, because we did a simulation test, and if the dam's design were faulty this would have registered in the simulation. Again, these are not intended as valid arguments, but as belonging to our repertoire of defeasible empirical reasoning about what is the case.

III

The Incompatibility Between the Suppositional View and the Propositional View. Now, conditional probability—the probability of C on the supposition that A—is not a measure of the probability of the truth of a proposition. There is no proposition X such that, necessarily, the probability that X is true is the conditional probability of C given A. If subjunctives are to be understood in terms of conditional probabilities, they are not to be understood in terms of truth conditions. For if they are to be understood in terms of truth conditions, you should believe a subjunctive to the extent that you think it is probable that it is true—that its truth conditions are satisfied.

Bennett (2003, pp. 254–6) claims that none of the arguments against truth conditions for indicative conditionals work for subjunctive conditionals because they all have at least one false premiss. He cites, *inter alia*, an argument of mine which includes the premiss that if you are certain that A or B without being certain that A, you must be certain that if $\neg A$, B, and points out that this is false for subjunctives. For example, I can be certain that either Oswald killed

[4] I owe this example to Arif Ahmed.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

Kennedy or someone other than Oswald killed Kennedy, without being certain that if Oswald hadn't killed Kennedy someone else would have. But Bennett misses the point that the conditional probabilities relevant to the assessment of subjunctive conditionals do not (typically) represent your present actual distribution of belief, but those of a hypothetical belief state in a different context, normally that of an earlier time, concerning, for example, whether someone else will kill Kennedy if Oswald doesn't. (The relevant translation of the above premiss would be: if you think it *was* certain that *A* or *B* *would* happen, but it was not certain that *A* would happen, you must be certain that if *A* hadn't happened *B* would have.)

The arguments Bennett discusses were aimed at indicative conditionals. They apply the general structural fact that a conditional probability does not measure the probability of the truth of a proposition to the case where the conditional represents your actual present state of conditional belief.

Here is another way of looking at the general structural fact. The fact that a conditional probability is not the probability of the truth of a proposition is in a sense the same structural fact as the fact that quantifiers like 'most' and 'almost all' in 'Most *A*s are *B*' and 'Almost all *A*s are *B*', or '90% of *A*s are *B*', unlike the standard treatment of the quantifiers 'all' and 'some' in 'All *A*s are *B*' and 'Some *A*s are *B*', are essentially binary, restricted quantifiers, in that they cannot be reduced to unary, unrestricted quantifiers: 'Most things (in the domain) are …' For probability statements can be reduced to, or at least modelled by, statements about proportions. Let me divide logical space into a finite number of (in my judgement) equiprobable bits, adequate for the problem at hand, i.e., every proposition I am concerned with is true throughout, or false throughout, any bit. For the sake of familiarity I shall call the bits 'worlds'; though they are not ultimate not-further-subdividable possibilities, they are divided finely enough for the project at hand. A proposition *B* is probable iff it is true in most of the worlds. It is almost certain iff it is true in almost all of the worlds. It is 90% probable iff it is true in 90% of the worlds. A proposition *B* is conditionally probable on the supposition that *A* iff most *A*-worlds are *B*-worlds; almost certain if almost all *A*-worlds are *B*-worlds; 90% probable if 90% of the *A*-worlds are *B*-worlds. If these were equivalent to statements about the probability of some proposition *X*, they would be equivalent to something of the form: in most worlds, *X* is true; most worlds are *X*-worlds; almost

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

all worlds are *X*-worlds; 90% of worlds are *X*-worlds, etc.; and we would have expressed the 'most' in 'Most *A*s are *B*' as a unary quantifier, which cannot be done.

In the general theory of quantifiers the first predicate is sometimes called the 'restrictor'. That is just what the antecedent of a conditional does: it restricts the claim that *C* to a hypothetical context in which the antecedent, *A*, is true.

IV

Arguments Against Truth Conditions for Counterfactuals. First, the relation between many central cases of 'wills' and 'woulds' is so close that it would be unreasonable to give a different account of each. I say 'If you touch that wire you will get a shock.' You don't touch it. I use my circuit-testing instrument to show you: 'You see, if you had touched it you would have got a shock.' Or, if the result is different: 'Funny, the power must be off. I was wrong. You wouldn't have got a shock if you had touched it.' A dog almost always, but not quite always, attacks and bites when strangers approach. I'm told 'It's very likely that you will be bitten if you approach.' I don't approach. Trusting my informant, I say 'It's very likely that I would have been bitten if I had approached.' Fred asks his doctor if he will be cured if he has the operation. The doctor says 'We can't be sure, but I'm pretty sure—about 90% sure that you will be cured if you have the operation.' Fred declines the operation, and the doctor, with no new relevant information, says 'It's very likely that he would have been cured if he had had the operation.' If the 'wills' are assessed by conditional probability, surely the 'woulds' are too. For easier arithmetical examples: it's 90% likely that you will get a ball with a black spot if you pick a red ball; it was 90% likely that you would have got a black spot if you had picked a red ball. Such pairs could be multiplied indefinitely.

The argument to follow applies to all accounts of truth conditions which construe a subjunctive conditional as some kind of strict conditional, involving universal quantification over some set of worlds or possibilities, or spelled out in terms of entailment from some premisses including the antecedent. I shall stick to the popular Lewis-style truth conditions (roughly, a subjunctive $A \rightarrow C$ is true iff *C* is true at all closest *A*-worlds (Lewis 1973, 1979)), though the same

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. CVIII, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

points can be made about Goodman-style truth conditions (Goodman 1965). It also applies to Lycan (2001), who says a conditional $A \rightarrow C$ is true iff all real and relevant A-possibilities are C-possibilities; and to Bennett (2003), who fine-tunes Lewis's account. Arguably in all of my examples above, and certainly in the last three, the counterfactuals would not come out as highly probable, but as known to be plain false on these truth conditions. Consider the dog that almost always bites when strangers approach. We can't tell the difference between the cases in which it does and those in which it doesn't. Either there is a bit of indeterminism in play, or it depends on some undetectable subtle feature of the manner of approach. It's not the case, and we take it not to be the case, that in *all* the relevant worlds in which I approached I was bitten. So the truth condition is not satisfied, and we believe it is not satisfied: we think it's certainly false that if you had approached you would have been bitten, according to the truth condition. Similarly for the doctor who thinks it 90% likely that I will be cured if I have the operation, and later considers whether I would have been cured if I had had the operation. Her uncertainty depends in part on the fine details of what might have happened in the operating theatre. She is certain that the Lewis truth condition does not obtain: that in not all relevant operation-worlds I am cured. Yet she thinks that it's 90% likely that I would have been cured if I had had the operation. And most obviously of all, the balls in the bag: it is certainly false that in all relevant worlds in which I pick a red ball, it has a black spot. But I say that it's 90% likely that you would have got a black spot if you had picked a red ball.

It might be objected that the probability goes in the consequent of the conditional itself. That is, it's just plain true that, for example, if she had had the operation, there would have been a 90% chance of being cured. I have two replies to this objection. First, in all of these examples, it is far from obvious that all relevant antecedent-worlds have the same probability of being a C-world, or even that in all relevant worlds the consequent has a high chance of being true. The doctor who believes it's 90% likely that I would have been cured need not believe that that figure would be right for every relevant world in which I go ahead with the operation. Indeed, it is compatible with her belief that she thinks some ways in which the operation could have gone would have had a very low chance of success. Thus the rendering 'In all close A-worlds, it is probable that C' is incorrect. Second, even if the first reply is inoperative, it *sounds* con-

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

tradictory to say 'It's certainly not the case that if she had had the operation she would have been cured; but if she had had the operation it is 90% likely that she would have been cured.' That is, there are not really two distinct natural ways of hearing these uncertain conditionals. Scope distinctions are a great philosophical tool, but we don't naturally hear two readings of these sentences.

(And I think we can see from these examples that these judgements of, say, 90% probability that *C* would have happened if *A* had, are not judgements that it is 90% probable that some fact obtains. What fact? Could God know whether it is true that I would have picked a spotted ball if I had picked a red ball?)

Put together the point about the close links between wills and woulds and the direct argument about how easy it is for counterfactuals to be false on the standard truth conditions. Consider the popular position that conditional probability is the measure of believability of an indicative conditional, but a Lewis-style account is needed for subjunctives. Those who hold this position must say that someone may be very confident that you will be cured if you have the operation, or that the dog will bite if you approach, or that you will get a black spot if you pick a red ball; but then, when the antecedent is found to be false, but there is no other change in their evidence, they must claim that the corresponding subjunctives are definitely false. This seems to me to be an unfortunate combination.

A consequence of the above phenomenon is that a very large number of the counterfactuals we accept and assert turn out to be false on the standard truth conditions. Either because of indeterminism, or because determinism is too fine-grained for our everyday antecedents, or because the concepts used in our everyday antecedents don't fit nicely into laws of nature (which play a large role in Lewis's account of closeness, and of course play a crucial role for Goodman), there will be the odd close world in which you strike the match and it doesn't light, let alone odd worlds which falsify counterfactuals about human behaviour: 'If you had asked me to do it I would have done so', 'If Fred had been in London he would have got in touch', and so on.[5] (It's important to see that this does not depend on inde-

[5] Unlike some writers on this theme, I would put less weight on the cases where the probabilities are astronomically high though less than one, which a Lewisian might arguably ignore, and more weight on cases where the probabilities are, say, around 90%, that is, significantly different from certainties; that is, the possibility of error cannot be ignored, yet one does not want to judge the conditional to be certainly false.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

terminism. Consider the balls-in-a-bag example as the easiest. Assume determinism. I didn't pick a red ball. So, given the laws and the past, I couldn't have picked a red ball. The supposition that I picked a red ball floats free of the past and the laws. So there is nothing in determinism to say exactly what my hand movements would have been. And not all pickings would have resulted in a black spot.)

These difficulties get some attention in Bennett's (2003) book. His first reaction he calls the near-miss proposal: a subjunctive conditional counts as true iff the consequent is true in almost all the relevant *A*-worlds. This is equivalent to saying that it's true iff the relevant conditional probability is sufficiently high.[6] This is both vague and context-dependent, but I don't object to that. Here are some objections, however. First, the proposal allows a conditional to be true which happens to have a true antecedent and false consequent. Bennett amends the account by adding that this is not the case. Second, suppose, just for the sake of argument, that the threshold for truth is around 99%. (I'm aware that it will be vague and context-dependent, but that does not affect the arguments to follow.) Then if you know that the relevant probability of *C* given *A* is 99%, (and you know that the antecedent is false, so the first amendment doesn't kick in), you know enough to be sure that the conditional is true, that is, sure that if *A* had been the case, *C* would have been the case (on this account). But you are not: you are just 99% sure that if *A* had been the case, *C* would have been the case. Third, if the relevant conditional probability misses the threshold by a small amount—say it is 90%—you should say the conditional is definitely false, and utterly reject 'If *A*, *C*'. But you don't: you think it is 90% likely that *C* would have happened if *A* had. Fourth, the proposal falls foul of the lottery paradox.[7] As one may put it, probabilities go down on conjunction, but truth-values don't! 'If *A*, *B*' and 'If *A*, *C*' should entail 'If *A*, then $B \wedge C$'. Suppose the balls in the bag are numbered 1 to 100. (I pick a balls-in-a-bag example just to get the structure right.) 'If you had picked a ball, it wouldn't have been number 1' and 'If you had picked a ball it wouldn't have been number 2' can both be true, but 'If you had picked a ball it wouldn't

[6] Igal Kvart (*ms.*), who like me treats counterfactuals in terms of conditional probabilities, also adopts something like Bennett's 'near miss' proposal: a counterfactual is true if the relevant conditional probability is above a certain context-dependent threshold.

[7] I owe this point to John Hawthorne (2005).

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

have been number 1 and it wouldn't have been number 2' is false. Fifth, consider conditionals of the form 'If you had tossed the coin *n* times you would have got at least one heads'. Whatever the threshold, it seems absurd to hold that as *n* increases there is some value for *n* at which such conditionals suddenly switch from false to true. Sixth: probabilities change. It may be above the threshold on Monday that if she were to have the operation on Friday she would survive, below the threshold on Tuesday, above again on Wednesday. So the conditional is true on Monday, false on Tuesday, true again on Wednesday. So 'If she had had the operation on Friday, she would have survived' gets different truth-values with reference to different times. This is perhaps not a knock-down objection, but we usually think of truth-values as more lasting features of our claims than probabilities.

Bennett (2003, p. 252) also considers what he calls a 'more radical proposal: drop truth'; and considers it favourably, which I think is right; but then goes on rather surprisingly to say that it doesn't matter very much, and does not narrow the gap between indicatives and subjunctives.

Well, it does matter somewhat: without truth, we can no longer think of validity in terms of preservation of truth; we no longer have a ready-made systematic theory of embedded conditionals. Bennett continues to write as though two counterfactuals with the same antecedent and contradictory consequents might both be false (see Bennett 2003, p. 256); but on the 'drop truth' proposal, no two such conditionals can each have a probability of less than 50%: their conditional probabilities sum to 1. And it does somewhat reduce the gap: both are assessed as conditional probabilities; we avoid the embarrassment of being almost certain that the dog will bite if we approach, but being forced to pronounce the corresponding counterfactual certainly false.

What Bennett means is that the careful fine-tuning of the notion of closeness which has occupied many chapters of his book is still needed, whether we go for truth or probability. There is something right about this. All theories of subjunctives—Goodman's, Lewis's, mine—share what is essentially the same problem, that of specifying what you hang on to and what you give up when you make a counterfactual supposition: suppose such-and-such had been the case; what do you hold constant? For Goodman this is the problem of cotenability, for Lewis it's the problem of closeness, for me, it's

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

the problem of which context shift is appropriate, which probability distribution is the appropriate one, given that it is not the one that represents your present state of belief. The probabilistic, suppositional view can be presented in such a way that it is a close relative of Lewis's. First we must specify the class of relevant *A*-worlds. Where Lewis says the counterfactual is true iff *C* is true in all of them, otherwise false, I say: take a probability distribution over them, and figure out how likely it is that we have a *C*-world, given that we have an *A*-world. We both have the problem of specifying the class of relevant worlds.

I shall say more about this soon. But I shall first make a brief remark about Stalnaker's truth conditions. Stalnaker does not treat the conditional as a kind of strict conditional. He says the conditional is true iff the consequent is true at *the* closest *A*-world. (This has been less popular than Lewis's approach.) Now when the antecedent is false, there is never a unique closest *A*-world. Think of all the different hand-movements you could have made if you had struck the match or picked a red ball. He adopts the technique of supervaluations to deal with this fact. So what we get is that the conditional is true iff the consequent is true at all permissible candidates for closest *A*-world, i.e. all closest *A*-worlds, false iff the consequent is false at all closest *A*-worlds; otherwise, the conditional is indeterminate. Now this is not so uncongenial to my way of thinking, which is not surprising, as Stalnaker's original aim was to find truth conditions compatible with the probabilistic account, and to extend the account to subjunctives. (Although this aim did not succeed, he still tries to approximate it closely.) My complaints, transposed to Stalnaker's account, are that vast numbers of subjunctive conditionals just get the verdict 'indeterminate', and this is not very helpful. Second, the probability of the *truth* of a conditional is, for Stalnaker, still the same as it is for Lewis, and all my problem cases turn out to be definitely not true. And we do not have a well-developed theory for how to think about how likely it is that if *A*, *C*, when it is almost certainly indeterminate. There have been some attempts to develop such a theory, by van Fraassen (1976), McGee (1985), Stalnaker and Jeffrey (1994), and myself once (1991), but all ran in to difficulties. I am inclined to think that if there were anything promising to be discovered along these lines it would have been discovered by now. But that judgement might be premature.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

V

The Relevant Worlds. When Lewis gave his criteria for closeness in 'Counterfactual Dependence and Time's Arrow' (1979), he did so for what he called the 'standard resolution of vagueness' of the relation of similarity between worlds. While I prefer the locution 'context-dependence' to vagueness, I agree with him that, in context, counterfactuals may violate the criteria he proposed, which should be seen, at best, as the default way of interpreting them when there are no special indications to the contrary.

Lewis's criteria vindicate this picture of how we standardly assess counterfactuals: consider those *A*-worlds which are exactly like the actual world until shortly before the antecedent time, and (if the antecedent is actually false) diverge from the actual world at an inconspicuous fork, and obey the laws of the actual world after the time of the fork. Ask whether the consequent is true in all such worlds. Once we have diverged from the actual world it is of 'little or no importance' whether the relevant worlds are approximately similar to the actual world in matters of particular fact (Lewis 1979, p. 48).

Suppose this picture is correct. Then my account would take that class of relevant worlds, take a probability distribution over them at the time of the fork, and consider how likely it is that *C* is true. If we make the simplifying assumption that all these worlds are equally likely, this would be to consider in what proportion of them *C* is true.

This picture often serves, but it is not invariably correct. *Pace* Lewis, it is sometimes of crucial importance that we consider *A*-worlds which share with the actual world some particular fact, concerning a time later than the antecedent time, and which could not have been predicted in advance. Here is an example I have discussed elsewhere (Edgington 2004): the car breaks down on the way to the airport and I miss my flight. 'If I had caught that plane I'd be half way to Paris by now', I remark to the mechanic, who has just been listening to the radio. 'You're wrong', he tells me. 'It crashed. If you had caught that plane you would be dead by now.'

Assume that the plane was brought down by a rare chance event, very unlikely in advance (so that in advance, the plane was no different in terms of safety from any normal plane); and assume that my presence or absence on the plane had no causal bearing on the crash; then, it seems, the mechanic's remark is correct. That is, the relevant worlds to consider are those in which the plane crashed.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

As there was a very low chance that the plane would crash, this hindsightful counterfactual does not seem to match a rational forward-looking judgement, 'If you catch the plane you will be killed'. True, but note that if someone made that forward-looking judgement in advance, albeit without good reason, and I miss the plane and it crashes, we would be inclined to say she was right! Although the best-informed and most reasonable opinion to have in advance is one which matches the objective chance, that does not mean that the objective chance gives the ultimate verdict on the conditional. Perhaps the chance that *C* if *A* was high, but it turns out that *A* and not *C*. Perhaps the chance is significantly different tomorrow that, for example, she will be cured if she has the operation. Or perhaps a chance event like the plane crash switches the verdict.

This example and others like it suggest that the conditional probability we are concerned to estimate, for counterfactuals (and in a sense the ultimate verdict on some forward-looking wills), is the chance, at a time when *A* still had some chance of coming about, of *C* given *A and* any relevant, causally independent, subsequent facts that have a causal bearing on *C*. You have the chance back then. Then you eliminate the 'no crash' possibilities and take a probability distribution over the remaining possibilities. It's still a conditional probability, but not one that represents a reasonable degree of conditional belief at the earlier time.

Why do we assess them this way? I argue elsewhere (Edgington 2004) that it is these hindsightful counterfactuals that help us to make inferences to true conclusions: 'She must have missed her plane', someone says, surprised to see me; 'if she had caught it she would be dead.'

Bennett also incorporates this kind of case into his refinement of Lewis's account of closeness. There are other refinements and complications. I turn now to the case for liberalism: pretty well any acceptable indicative conditional can 'go counterfactual' in a suitable context. Here is an extreme example (borrowed and adapted from Grice 1989), which concerns a very minimal ground for an indicative. If the shift to the counterfactual is permissible here, it looks as if it is permissible, in a suitable context, for any indicative. There is a treasure hunt. The organizer tells me 'I'll give you a hint: it's either in the attic or the garden.' Trusting the speaker, I think 'If it's not in the attic it's in the garden.' We are competing in pairs: I go to the attic and tip off my partner to search the garden. I discover the treasure. 'Why did you

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

tell me to go to the garden?' she asks. 'Because if it hadn't been in the attic it would have been in the garden: that's (what I inferred from) what I was told.' That doesn't sound wrong in the context.

Or consider: 'Why did you hold Smith for questioning?' 'Because we knew the crime was committed by either Jones or Smith—if it hadn't been Jones, it would have been Smith.' There's also a nice example of van Fraassen's (1981): the conjuror holds up a penny and claims he got it from the boy's pocket. 'That didn't come from my pocket', says the boy. 'All the coins in my pocket are silver. If that had come from my pocket, it would have been a silver coin.'

These examples are of no great intrinsic interest, but they are not semantically defective. This takes off some of the pressure to find *the* account of relevance or closeness. It also allows us to make sense, in context, of 'far out' counterfactuals which do not easily fit the standard pattern outlined above. Nevertheless, I share Lewis's view that there is a default, most context-free way of assessing counterfactuals, which feeds into our inferential practices in important ways.

VI

The Relocation Thesis. I shall make a few remarks on the 'relocation thesis', the thesis that wills and woulds are one kind of conditional, the plain past/present indicatives another. V. H. Dudman (1984a, 1984b) was not the first to stress the close relation between wills and woulds (which I think is correct) but he was one of the first to draw the conclusion that there are two kinds of conditionals: the wills and woulds are one kind, the plain past and present tense indicatives are another kind.[8] This thesis has been quite influential, adopted by, *inter alia*, Bennett (1988) (who subsequently dropped it), Woods (1997), Smiley (1984), and Mellor (1993). Often in the philosophical literature the wills and woulds are treated as something like 'causal conditionals', the others as 'evidential conditionals'. I'm against splitting the traditional class of indicative conditionals in this way.

In the traditional class of indicative conditionals, we have a declarative sentence suitable for making a statement, be it about the past, present or future (or indeed timeless), to which a conditional

[8] Gibbard (1981), independently, has a similar thesis but thinks the 'wills' are ambiguous, and can function as either kind of conditional.

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

clause is attached, expressing a judgement not categorically but in the context of a supposition; that is, they do essentially the same sort of thing. Ramsey's thesis is plausible for all this class: you are confident in a conditional to the extent that you have a high degree of belief in the consequent on the supposition of the antecedent. Naturally, our grounds tend to be different for statements about the future and statements about the past, and a common and important sort of ground for the 'wills' is evidence for the claim that *A*, if it happens, will cause it to be the case that *C*. But first, this sort of ground can equally apply to conditionals about the past—'If she touched that, she got a shock'; and second, it is not the only kind of ground for those about the future: I know the boss told one of his assistants to meet me at the station, but I don't know which; so if Bob doesn't come, Ann will come. Even for those that cry out for a causal reading, one can tell more or less bizarre non-standard stories. Here is a striking one from Bennett (2003, pp. 343–4): 'If it rains tomorrow, the roads will be slippery.' But I don't mean that rain will make the roads slippery: the roads are very well constructed and not made slippery by rain. I've just received a leaflet from the council which (a) includes a weather forecast predicting rain; and (b) says they intend to oil the roads tomorrow, warning that this will make the roads slippery. It doesn't look as if it's going to rain, but the council has a first-rate weather forecaster. However, there is some reason to suspect that the leaflet may be a hoax and not genuine. If it rains, that will be evidence that it is genuine, and hence that they will oil the roads, and hence that the roads will be slippery. Of course, one would mislead by making that conditional remark without warning that the most obvious ground is not the operative one. But that is pragmatics. No conditional that does not explicitly use causal language like 'produce' or 'make', 'result' or 'outcome', forces a causal reading, though of course it is very often rightly presumed to be asserted on causal grounds. 'If *A* happens, *B* will happen, but *A* won't cause *B* to happen' is never contradictory.

VII

Concluding Remarks. None of the main theories of conditionals is incoherent. All are possible ways in which speakers and thinkers could use 'if'. It is an empirical question which theory fits our prac-

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

tice best. Why do philosophers get worked up about it? Why don't we just leave the matter to be settled by questionnaires, or the empirical work of linguists and cognitive psychologists? (And indeed there is much work in this field; see Evans and Over 2004.) It is not just an empirical question for philosophers. It is a normative question. We have here an immensely valuable form of thought, without which our thinking would be immeasurably diminished. And we want the theory that best explains why conditionals matter so much to us. As I have said before, the truth-functional theory of indicative deprives us of the ability to distinguish between believable and unbelievable conditionals whose antecedent we think is unlikely to be true. We would be intellectually impoverished if we used 'if' that way. And, I have argued today, a lot of theories of subjunctive conditionals have the consequence that almost all but the most trivial conditionals of this form are knowably false; and this would have a disastrous effect on the use we make of these conditionals. We get worked up because we have the inkling that there is an essential form of thought here, which serves important purposes, and we are after the nature of conditional thinking—an account of how and why this form of thought serves important purposes.

If it is taken for granted that we are seeking truth conditions—a theory of the circumstances in which a conditional judgement is true—one is apt to ignore uncertain conditional judgements, and test one's theory on examples in which someone would claim to know that if *A*, *C*. You need not deny that there are uncertain conditional judgements, but they are not your special business—they are the business of a general theory of uncertainty about what is true. They are not in focus. But when we do focus on uncertain conditional judgements we find that truth-conditional theories have bad consequences—they either underestimate or overestimate that uncertainty. And we find we have ready made a valuable concept for understanding uncertain conditional judgements, the concept of conditional probability.

School of Philosophy
Birkbeck College
University of London
Malet Street
London WC1E 7HX
UK

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. CVIII, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

References

Adams, E. W. 1975: *The Logic of Conditionals* (Dordrecht: Reidel).
——1993: 'On the Rightness of Certain Counterfactuals', *Pacific Philosophical Quarterly*, 74, pp. 1–10.
Appiah, Anthony 1985: *Assertion and Conditionals* (Cambridge: Cambridge University Press).
Bennett, Jonathan 1988: 'Farewell to the Phlogiston Theory of Conditionals', *Mind*, 97, pp. 509–27.
——2003: *A Philosophical Guide to Conditionals* (Oxford: Oxford University Press).
Dudman, V. H. 1984a: 'Conditional Interpretations of *If*-sentences', *Australian Journal of Linguistics*, 4, pp. 143–204.
——1984b: 'Parsing "If"-sentences', *Analysis*, 44, pp. 145–53.
Edgington, Dorothy 1991: 'The Mystery of the Missing Matter of Fact', *Proceedings of the Aristotelian Society Supplementary Volume* 65, pp. 185–209.
——2004: 'Counterfactuals and the Benefit of Hindsight', in P. Dowe and P. Noordhof (eds.), *Chance and Causation* (London: Routledge, pp. 13–28).
Evans, Jonathan and David Over 2004: *If* (Oxford: Oxford University Press).
Gibbard, Allan 1981: 'Two Recent Theories of Conditionals', in W. Harper, R. Stalnaker and C. Pearce (eds.), *Ifs* (Dordrecht: Rediel, pp. 211–47).
Goodman, Nelson 1965: *Fact, Fiction and Forecast* (Indianapolis: Bobbs-Merrill.
Grice, H. P. 1989: 'Logic and Conversation', in his *Studies in the Way of Words* (Cambridge, MA: Harvard University Press).
Hawthorne, John 2005: 'Chance and Counterfactuals', *Philosophy and Phenomenological Research*, 70, pp. 396–405.
Jeffrey, R. 1991: 'Matter of Fact Conditionals', *Proceedings of the Aristotelian Society Supplementary Volume* 65, pp. 161–83.
Kvart, Igal *ms*.: 'The Causal-process-chance-based Analysis of Counterfactuals'.
Lewis, David 1973: *Counterfactuals* (Oxford: Blackwell).
——1976: 'Probabilities of Conditionals and Conditional Probabilities', reprinted in Lewis 1986, pp. 133–56.
——1979: 'Counterfactual Dependence and Time's Arrow', reprinted in Lewis 1986, pp. 32–66. Page references to this volume.
——1986: *Philosophical Papers II* (Oxford: Oxford University Press).
Lowe, E. J. 1995: 'The Truth About Counterfactuals', *Philosophical Quarterly*, 45, pp. 41–59.
Lycan, William 2001: *Real Conditionals* (Oxford: Oxford University Press).

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x

——2005: Review of Bennett 2003, *Mind*, 114, pp. 116–19.
McGee, Vann 1989: 'Conditional Probabilities and Compounds of Conditionals', *Philosophical Review*, 98, pp. 462–71.
Mellor, D. H. 1993: 'How to Believe a Conditional', *Journal of Philosophy*, 90 (5), pp. 233–48.
Smiley, Timothy 1984: 'Hunter on Conditionals', *Proceedings of the Aristotelian Society*, 84, pp. 113–22.
Stalnaker, Robert 2005: 'Conditional Propositions and Conditional Assertions', in *New Work on Modality: MIT Working Papers in Linguistics and Philosophy*, vol. 51.
Stalnaker, Robert and Richard Jeffrey 1994: 'Conditionals as Random Variables', in E. Eells and B. Skyrms (eds.), *Probability and Conditionals* (Cambridge: Cambridge University Press, pp. 31–46).
Strawson, P. F. 1952: *Introduction to Logical Theory* (London: Methuen).
——1986: '"If" and "⊃"', in R. Grandy and R. Warner (eds.), *Philosophical Grounds of Rationality: Intentions, Categories, Ends* (Oxford: Oxford University Press, pp. 229–42).
van Fraassen, Bas 1976: 'Probabilities and Conditionals', in W. Harper and C. Hooker (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, vol. 1 (Dordrecht: Reidel).
——1981: 'Essences and Laws of Nature', in R. Healey (ed.), *Reduction Time and Reality* (Cambridge: Cambridge University Press, pp. 189–200).
Woods, Michael 1997: *Conditionals* (Oxford: Oxford University Press).

©2008 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. cviii, *Part* 1
doi: 10.1111/j.1467-9264.2008.00233.x