

14 Ramsey's Tests

1

The so-called 'Ramsey Test', as a way of understanding conditionals, has been discussed very extensively over the last few decades. Robert Stalnaker formulated the 'Ramsey Test' as follows (Stalnaker 1969, 102):

First add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the belief in the antecedent); finally, consider whether or not the consequent is then true.

Stalnaker produced an axiomatic system for conditionals on this basis which replaced the 'stocks of beliefs' with possible worlds. According to his theory, a conditional 'if A then B' is true in a possible world i if and only if B is true in possible world $f_i(A)$, where the latter function is a selection function satisfying certain conditions. Peter Gärdenfors formulated the 'Ramsey Test' somewhat differently: 'Accept a proposition of the form 'if A then C' in a state of belief K if and only if the minimal change of K needed to accept A also requires accepting C.' (Gärdenfors 1986, 81). Thus, the formal incorporation of the 'Ramsey Test' in Gärdenfors' work on belief revision brings in, for instance, the axiom

if $K \neq K_{\perp}$, and $K_A = K_{\perp}$, then $\vdash \neg A$,

where K_A is the belief set adjusted to accommodate A, and with

$A > C \in K$ iff $C \in K_A$.

Now there might be arguments about how well either of these writers have formalised the notions of 'adjustment' and 'minimal change'. Stalnaker's idea was that the chosen world

would be 'nearest', in some sense, to the world i , although that sense was not exactly specified. And Gärdenfors' axiom hardly determines the new belief set, since it merely means that if K and A are both consistent then K_A should be consistent, and there may be many ways to achieve this. But Ramsey's historical account is significantly different from each of these ideas, even to the extent that, in key cases, it did not involve *any* adjustment, or change of the belief set *at all*. Certainly Ramsey's general description of the conditional is as follows (Ramsey 1978, 144):

'If p , then q ' can in no sense be true unless the material implication $p \supset q$ is true; but it generally means that $p \supset q$ is not only true but deducible or discoverable in some particular way not explicitly stated. This is always evident when 'If p then q ' or 'Because p , q ' (*because* is merely a variant on *if*, when p is known to be true) is thought worth stating even when it is already known either that p is false or that q is true. In general we can say with Mill that 'If p then q ' means that q is inferable from p , that is, of course, from p together with certain facts and laws not stated but in some way indicated by the context.

But, more specifically, Ramsey's prime suggestion was (Ramsey 1978, 143):

If two people are arguing 'if p will q ?' and are both in doubt as to p , they are adding p hypothetically to their stock of knowledge and arguing on that basis about q : so that, in a sense, 'If p , q ' and 'If p , $\neg q$ ' are contradictory. We can say they are fixing their degrees of belief in q , given p . If p turns out false, these degrees of belief are rendered *void*. If either party believes $\neg p$ for certain, the question ceases to mean anything to him except as a question about what follows from certain laws or hypotheses.

So there are substantial differences even here with how Ramsey has been taken by the above writers. Stalnaker formulated his version of all this thinking that Ramsey's suggestion about adding p hypothetically covered only the case in which one has no opinion about the truth value of the

antecedent (Stalnaker 1969, 101). But Ramsey's suggestion, at this place, evidently covered also the case where the antecedent is disbelieved. Moreover, he did not then envisage maintaining consistency, in order to incorporate p into the belief set – and see what would then be true. The question, according to Ramsey, then 'ceases to mean anything' unless it is construed a different way, which is to say, first of all, that the indicatives 'if p , q ' and 'if p , $\neg q$ ' are then both true, for Ramsey. And that naturally arises upon simply adding p to the belief set, since it becomes inconsistent in that case.

But more important, if the question was turned into one about what followed from 'certain laws and hypotheses', it was not an adjustment or change of the belief set which was needed, according to Ramsey. Subjunctives have been sometimes called 'backtrackers', but Ramsey's account showed them not to be 'tracking back'. Instead, they required a certain *causal* backing: 'Causality' is even in the title of the paper from which the above quotations come. Ramsey amplified his account of the case of disbelief, and its further relation to 'certain laws and hypotheses' in this way (Ramsey 1978, 145):

One class of cases is particularly important, namely those in which, as we say, our 'if' gives us not only a *ratio cognoscendi* but also a *ratio essendi*. In this case, which is e.g. the normal one when we say 'If p had happened, q would have happened', $p \supset q$ must follow from a hypothetical $(x)(\Phi x \supset \Psi x)$ and facts r , $pr \supset q$ being an instance of $\Phi x \supset \Psi x$, and q describing events not earlier than any of those described in pr . A variable hypothetical of this sort we call a *causal law*.

Thus Ramsey was well aware of the difference between indicative and subjunctive conditionals, as his distinction between *ratio cognoscendi* and *ratio essendi* indicates. In addition, he says (Ramsey 1978, 145):

Corresponding to the kind of laws or facts intended we get various subtle syntactical variations. For instance, 'If he was there, he must have

voted for it (for it was passed unanimously), but if he had been there he would have voted against it (such being his nature)'.

This shows that, according to Ramsey, the indicative and subjunctive conditionals are supported in quite different ways. The indicative 'if p , q ' means that q is inferable from p given the background information, which means it is true if and only if $K, p \vdash q$, where K is that background information. But the subjunctive form 'If p had happened, then q would have happened' is backed in a much more specific way by the evidence, according to Ramsey. For, in the first place, $(x)(\Phi x \supset \Psi x)$ has to be inferable from K , and the further inference from this, namely $pr \supset q$ must contain an r which is also part of K . Hence the subjunctive is not obtained by adding a certain antecedent to the belief set, let alone by adding something to the belief set and adjusting it to be consistent, or whatever.

That seems quite natural. Consider, for instance, the following: 'He does not have AIDS, but if he had AIDS, he would have previously been infected with the HIV virus.' Assessing this does not require that we shift from one belief set to another: from a belief set where the subject does not have AIDS to one, after the 'but', where he does have, and in which 'he was previously infected by the HIV virus' non-contradictorily occurs. On a natural understanding of belief sets, the same belief set which includes the fact that someone does not have AIDS clearly may also include the relevant causal law – which is what *pre-determines* that most possible relevant cases verify the indicative conditional 'If he does have AIDS, he was previously infected by the HIV virus', whether or not 'he has/does not have AIDS' is then true. We do not (indeed, how could we) inspect those possible cases to verify the law; it is the law itself which, in part, determines which cases are possible. So what is first required, with the subjunctive, is some statement form which can express such a causal relation. Ramsey's historical account of subjunctive

conditionals is thus very much at variance with subsequent versions of it – although, as we shall see, it has considerable affinities with the causal account of subjunctive conditionals recently given by Judea Pearl (Pearl 2000).

2

It is already clear, however, that Ramsey's historical account of conditionals is *better* than subsequent versions of it. How it relates to Pearl's work we shall see later, but the principles determining its superiority over Stalnaker and Gärdenfors are already present in Ramsey. This is for several reasons: first the consequent relation to probability of *just subjunctive* conditionals; for not all conditionals were probabilistic, according to Ramsey.

Ramsey's point about when the pure indicative conditional is assertable, we can now see, is no more than an assertion of the Deduction Theorem. It is to the effect that

$$K \mid - p \supset q \text{ iff } K, p \mid - q,$$

where ' \supset ' is the ordinary material conditional. Ramsey's remarks about the case where there is doubt about the antecedent are obviously in tune with this. But also his remarks about the case where the antecedent is disbelieved support this interpretation. For then, what modern followers clearly have not respected, Ramsey still did take the indicative conditional to be material (and see Slater 1996 for a further defence of this). As above, Ramsey took the indicatives 'if p , q ' and 'if p , $\neg q$ ' to be both true, when the antecedent is untrue, with meaningfulness, i.e. a non-contradictory consequent, only being recoverable in a different, if allied case, where the indicative conditional is replaced by a subjunctive version of itself, related to some quantified version of itself. And it is, of course, only in connection with such quantification that

probabilities can be assigned, since these are measured in fractions, i.e. fractions of alternative cases.

It is here, though, that we might want to depart from some of the extant specifics of Ramsey's account, even while we acknowledged it was a great improvement on recent formulations. One modification which might seem to be needed, for instance, relates to the character of Ramsey's universal formula. Similar quantifications, nowadays, are commonly taken over different possible worlds, rather than different cases in this world, as is more appropriate to an empiricist approach like that of Ramsey. Within each world (rather than for each case), we will then want to say, the material conditional will or will not be true, so over many worlds (rather than many cases) we may be able to make the more general, subjunctive remark.

But if q is then going to be inferable from p only in certain circumstances r , the relation just between (generalisations of) p and q may itself not be strictly universal. So, first, a probabilistic expression, more in tune with Ramsey's wider thinking on these matters, best expresses this:

$\text{pr}(\text{if } p \text{ were the case then } q \text{ would be the case}) = \text{pr}(q/p)$.
Indeed, while this is often called 'Adams' Hypothesis', it has also been attributed directly to Ramsey himself – as well as Stalnaker. But, as a result, we can proceed, first, to investigate just this formalistic expression, and leave it until later to arbitrate between giving the probability an interpretation in terms of possible worlds, instead of, what is certainly more historically accurate as an account of Ramsey, namely a frequentist interpretation in terms of proportions of cases.

3

It is surprising that the general power of probabilistic accounts of (subjunctive) conditionals is not universally appreciated.

The major source of confusion must be the influence of Lewis' so-called 'triviality results' (Lewis 1976), since they have been taken to negate Ramsey's Hypothesis. Later I shall correct this, but still Lewis' triviality results, historically, seem to have turned much of contemporary thought away from the natural, probabilistic way of looking at things.

Forgetting the power of Ramsey's account in particular seems to have hindered seeing what is surely now the most patent and obvious resolution of the famous situation described by Allan Gibbard (Gibbard 1981):

Zack, knowing Stone's poker hand is quite good, and having told Pete of its contents (but unaware of Pete's hand) confides 'If Pete called, he won'. Jack, however, has seen both hands, and knows Pete's hand is rather low, so he is in a position to say 'If Pete called, he lost'. But if we accept both conditionals we are obliged to conclude Pete folded, which is not necessarily the case at all.

One notes how steeped this example is in epistemic notions, like states of knowledge. But the normal inclination to read such epistemic situations probabilistically has evidently been suppressed, since no subsequent discussion of this case, to the present writer's knowledge (see, e.g. Stalnaker 1987, Pendelbury 1989, Jackson 1990, 1991, Edgington 1995), has seen the immediate, probabilistic way out.

The point is that Zack's evidence for his conditional is not as conclusive as Jack's is for his, and so, while Jack is entitled to assert his conditional without qualification, Zack, strictly, can only say 'Most probably, if Pete called, he won'. Pendelbury (Pendelbury 1989, 182) considers the possibility that one conditional is true, and the other false, and so maybe partly senses the crucial point at issue. But he does not consider more measured evaluations of conditionals, like 'most probably true' – which is the relevant one in this case. For there is no conflict between its being most probable that something happened,

and its not, in fact, happening. Hence there is no basis to conclude, against the possible fact, that Pete folded.

I have shown elsewhere how inattention to such qualifications as 'probably' has distorted the debate over conditionals (Slater 1992b), and maybe that has again arisen in this case. But that is not the end of the lessons which come from returning to a probabilistic understanding of these things. Other lessons, for instance, bear closely on Belief Revision theory, as illustrated in the piece from Gärdenfors at the start. Thus, amongst other things, Adjunction fails for probabilistic belief, and so we do not have necessary deductive closure of the belief set. Also, with a probabilistic account of belief, as Ramsey himself showed, we have a more natural account of the human mind (Slater 1993b). There might seem to be a difficulty with belief revision using conditionalisation in the case where some new fact contradicts the old belief set. Revision of belief sets in such circumstances is allowed in Belief Revision theory. But since the fact's anterior probability is zero, conditionalisation means that the probability of it, were it to be included, would also be zero. So the supposition is itself a contradictory one, since the supposed new 'fact' would be judged a falsehood on the basis of the previous belief set, and so no revision to accommodate it would be envisioned. The point, of course, is connected with just that aspect of Ramsey's remarks which has been overlooked: the basis for a counterfactual conditional comes from within the existing belief set, not some revision of it.

4

The major trouble with any sort of probabilistic account of subjunctive conditionals, however, is that, as before, it is usually thought that Lewis' triviality results block them. Lewis showed, in several theorems, that there is no way to interpret a conditional connective so that the probability of a conditional,

in general, is the associated conditional probability. What still remains true, however, is that, for a certain subclass of conditionals – the subjunctive conditionals, of course – their probability is a conditional probability. I will provide a formal proof of this later, but the basis for the re-orientation is also found in Pearl. Pearl does not discuss Lewis' theorems explicitly, but he clearly realises what is needed to avoid them, when making a point rather similar to Ramsey's on facts and laws. The conditionals whose probability is a conditional probability are those associated with laws. Pearl says (Pearl 2000, 224-5):

Facts are expressed in ordinary propositions and hence can obtain probability values and can be conditioned on; laws, on the other hand, are expressed as conditional probability sentences (e.g. $P(\text{accident/careless driving}) = \text{high}$) and hence should not be assigned probabilities and cannot be conditioned on. It is because of this tradition that probabilists have always attributed non-propositional character to conditional sentences (e.g. birds fly), refused to allow nested conditionals ... and insisted on interpreting one's confidence in a conditional sentence as a conditional probability judgement ...

Making the distinction between laws and facts, indeed, does depend on introducing items which are not propositions in the normal way, and departing, in the ways Pearl itemises, from the classical logic of propositions. But the specific grammatical process which enables us to do this Pearl does not detail: it requires predicate logic's extension by means of epsilon terms, as I have shown in several places before (for instance, Slater 2000b, 315-6). There is a close connection, nevertheless, between the relevant epsilon terms and the novelty Pearl does introduce, namely his '*do*-operator', as we shall see. So let us first have clear some further details of Pearl's account, before we see just what the logical mechanism is which allows the probability of certain conditionals to be a conditional probability, despite Lewis' results.

Pearl develops in fine detail a probabilistic, agency account of causation, which is the sort of account of causation commonly attributed to Ramsey (Menzies and Price 1993, 187). The way agency, and a form of probability enter Pearl's 'interventionist' account can be seen from the following (Pearl 2000, p351):

If we wish to find the chance that it rained, given that we see the grass wet, we can express our question in a formal sentence written like that: $P(\text{Rain} \mid \text{Wet})$, to be read: the probability of Rain, given Wet. The vertical bar stands for the phrase: 'given that we see' ... But suppose we ask a different question: 'What is the chance it rained if we *make* the grass wet?' We cannot even express our query in the syntax of [standard] probability, because the vertical bar is already taken to mean 'given that I see'. We can invent a new symbol *do*, and each time we see a *do* after the bar we read it *given that we do* ...

But the parallel with Ramsey does not stop there, since the specific way in which Pearl handles subjunctive conditionals also has close affinities with Ramsey's treatment, since, of course, both are based on causality. When contrasting Lewis' 'similarity' account of counterfactuals with his own 'structural' account, Pearl says (Pearl 2000, 238-239):

Implicit in [Lewis'] proposal lies a claim that counterfactual expressions are less ambiguous to our mind than causal expressions. Why else would the expression 'B would be false if it were not for A' be considered an *explication* of 'A caused B', and not the other way around, unless we could discern the truth of the former with greater certitude than that of the latter? Taken literally, discerning the truth of counterfactuals requires generating and examining possible alternatives to the actual situation as well as testing whether certain propositions hold in those alternatives – a mental task of nonnegligible proportions ... Such difficulties do not enter the structural account. In contrast with Lewis's theory, counterfactuals are not based on an abstract notion of similarity among hypothetical worlds; instead, they rest directly on the mechanisms (or 'laws', to be fancy) that produce those worlds and on the invariant properties of those mechanisms.

A finer analysis of the similarity between Ramsey and Pearl will appear later. I shall first give a grammatical characterisation, in epsilon terms, of subjunctive conditionals, show their intimate relation to probability, and also their relation to Pearl's *do*-operator.

5

Epsilon terms, we know, are defined by the equation

$$(\exists x)Fx \equiv F\epsilon xFx,$$

where ' ϵxFx ' may be read 'the first F '. But let us write 'it would be true, in world i , that p ' (i.e. what is otherwise written ' $V(p, i)=1$ ') as ' Wip '. On this basis we get, for instance,

$$(\exists i)Wip \equiv W(\epsilon iWip)p,$$

where the brackets are inserted just for ease of reading, and the epsilon term ' $\epsilon iWip$ ' selects some world which is a p -world, if there are any – and any world at all, if there are not. As a result, it is closely allied with Pearl's *do*-operator, as we shall see, but it also gives a natural representation for the anaphoric phrase 'that case' which occurs in subjunctive forms like

If there were chickens, in that case there would be eggs,

i.e.

$$(\exists i)Wic \supset W(\epsilon iWic)e.$$

The cross reference to the antecedent world is supplied by the ' $\epsilon iWic$ ' in the consequent here, since that is also the epsilon term hidden in the antecedent. Moreover the natural conditional grammar is preserved, and there is no difficulty in still using ' \supset ' for 'if'. The subjunctive nature of the conditional is now put entirely into the content of the antecedent and consequent, as in natural language.

The distinctive nature of conditionals of this form is not widely recognised, but Brian Ellis pointed out one very significant fact about them when dealing with

If there is a man of thirty, he is married.

Ellis remarked that such an expression does not have a simple truth value, which is another way of putting Pearl's point about them not being straightforwardly propositional (Ellis 1966, 170):

[it] is neither true nor false. To say that it is true is to imply that all men of thirty are married. To say that it is false is to imply that no men of thirty are married. The only thing we can say is that it is probable that a man of thirty is married.

Ellis, as a result, looked forward to the day when the probability and predicate calculi would be united into a single calculus, and the epsilon calculus in fact enables us to do just that. For, while subordinate propositions of the required kind are not available within the expressive limits of the predicate calculus they are available within the expressive limits of the epsilon calculus – and we can add probability and other operators to it (see, for instance, Purdy 1994). Conditionals like the above are 'indefinite propositions' in Ellis' terms (Ellis 1966, 168), and they contain terms (here epsilon terms) which have what Ellis called 'indefinite reference'.

The specific formal advantage of the epsilon analysis is that, although epsilon terms are like Stalnaker's selection functions in a broad way, their definition by means of the equation above means we get immediately what is required, i.e. we get conditionals which obey 'Ramsey's (Adams', Stalnaker's) Hypothesis'. For, if there can be chickens in the world, the probability of

$$\neg(\exists i)Wic \vee W(\epsilon iWic)e,$$

is just the probability of the second disjunct, i.e. the chance that the chosen c-world should be an e-world. But this is just $\text{pr}(c.e)/\text{pr}(c)$, i.e. the conditional probability $\text{pr}(e/c)$. On the other hand, if there cannot be chickens in the world, the probability of the disjunction is 1, which we can take to be the conditional probability (by stipulation) in that case. Stalnaker assumes there is an 'absurd world' in which everything is true

to handle this side of the matter, but, with the probability of the first disjunct being 1 in this case, the result is automatic.

The reason Lewis' triviality results do not apply is because, as Pearl indicated, subjunctive conditionals, on this understanding, cannot be further conditionalised, i.e. they are absolute. Lewis conditionalised the proposed conditional whose probability was a conditional probability to get his results. But he merely presumed that this could be done, and so his results only hold when indeed it can be done. The absoluteness of the present conditionals follows from the fact that they are laws in Ramsey's, and Pearl's sense, so they pre-determine possibilities rather than are dependent on them.

Formalistically this is connected with another feature of the epsilon account which separates it from Stalnaker's: the non-relativity of the selection function to worlds (as with Stalnaker's $f_i(A)$). It might seem that such a relativity had to be allowed for, because of examples like: 'If Di had gone to the party, Charles would not have gone – though that need not have been the case, if they had not separated.' But what this is saying is that $\text{pr}(\neg c/d)$ may be near to 1 while $\text{pr}(\neg c/d.\neg s)$ is not, i.e. it is an example of the failure of Antecedent Restriction, as on all accounts. The probabilistic account does not deal in 'nearest worlds', but most probable outcomes, which makes it quite distinct in another way from Stalnaker's account. For p may be highly probable, and likewise q, without p.q being highly probable (i.e. as mentioned before, Adjunction fails). So what is highly likely does not necessarily form into a world, let alone a 'nearest' one.

So what is the relation between the epsilon account, and Pearl's account in terms of the *do*-operator? Pearl asks us to consider, in order to evaluate counterfactuals, not the antecedent 'given that we see', but the antecedent 'given that we do' in our probability calculations. But ' $\epsilon iWic$ ' selects some world which is a c-world, if this is possible, so the antecedent in the epsilon conditional, ' $W(\epsilon iWic)c$ ', encapsulates the action

of making a c-world. Of course, in the case discussed above, making it so that there are chickens results inevitably in there being eggs, validating the counterfactual in question. The point also settles the issue left over from before, about the appropriateness of talking of 'possible worlds' in connection with Ramsey's account – which is related to the epistemological difficulties Pearl pointed out in Lewis' account. For, given the link with action, what is of concern is what is 'possibly the world' as a result of such actions – not a 'possible world' as some other reality, in the sense of Lewis. Interventions in the form of scientific experiments alter this world, and so build up the requisite number of empirical cases to which the frequency definition of probability can be applied. No other cases but cases in this world are thus involved; indeed no other could be the proper basis for knowledge of scientific laws. But what the pure Empiricist forgets is that, in addition to being observers, we are also agents in the world, and so what is possibly in it is not given independently of us, and our actions. 'In a sense my present action is an ultimate and the only ultimate contingency' (Ramsey 1978, 146).

6

It remains to contrast some more detailed aspects of Pearl's account with the epsilon version of Ramsey's. Although there clearly is a major similarity in orientation between them, there are also some significant differences.

The most obvious difference, of course, is that, as Ramsey wanted, the epsilon account gives us an explicit conditional whose probability is a conditional probability. But this feature leads immediately to several specific results regarding the full logic of subjunctive conditionals. Indeed, just on the above basis, i.e. the definition of epsilon terms, there is available, from first principles, a very extensive account. I have shown,

for instance, that what are standardly called the failure of Contraposition, Antecedent Restriction, and Hypothetical Syllogism immediately follow (Slater 1988a, and 1988d, 133-43), even though the material connective is still involved. Thus there is no equivalence between

$$(\exists i)Wic \supset W(\exists iWic)e,$$

and

$$(\exists i)Wi\neg e \supset W(\exists iWi\neg e)\neg c.$$

But what makes the equivalence fail is not some feature of some new connective, merely the specific contents of the antecedents and consequents in question.

For further illustration I will now present some facts about the case where all the probabilities are 1, although it must be stressed that this is just one, special case. It is significant to consider it in the present connection, however, since Pearl gives a comparison of his system with Lewis' in this case (Pearl 2000, 240-2), and so a comparison between the epsilon system and Lewis' system, as well as Pearl's can be obtained at the same time. Pearl's system is equivalent to Lewis' in this case, except in one respect.

For Lewis, all counterfactuals are exceptionless, of course, since his formula, here written ' $A \Box \rightarrow B$ ', stands for 'In all closest worlds where A holds, B holds as well'. What happens, therefore, when we replace ' $A \Box \rightarrow B$ ' with ' $\text{pr}(B/A)=1$ ', in Lewis' axioms (Lewis 1981, 80)? As readers may check for themselves, the only axiom which is at all questionable is

$$A.B \Rightarrow A \Box \rightarrow B.$$

But this has been questioned before. Certainly given A.B it does not follow that had B been false, A would have been false, i.e. that $\text{pr}(\neg A/\neg B)=1$, but still it does follow that $\text{pr}(B/A)=1$. So understanding the fact that Contraposition fails (in the above way) is crucial to recognising this axiom as valid. And Pearl's system endorses it, along with all the others. The only axiom in Pearl's system not endorsed by Lewis, by contrast, is

'reversibility' (Pearl 2000, 229). This amounts to, in probabilistic terms,

$$[\text{pr}(y/wx)=1, \text{pr}(w/yx)=1] \Rightarrow \text{pr}(y/x)=1,$$

which is not valid, as is readily seen by taking $y=w$. So it is not endorsed by the epsilon analysis either.

There may be further significant differences of quite a different kind, between Ramsey's system and Pearl's, since, like Lewis, Pearl treats causation as a primary quality – that is one reason why his is a 'structural' account. But Menzies and Price, for instance, have argued that Ramsey's set of ideas is best associated with seeing causation as a secondary property, i.e a property intimately related not to the world as such, but to our experience of it. Whatever holds in this more philosophical area, however, formally the epsilon version of Ramsey's ideas is surely more principled, and persuasive, as a consequence of the above comparisons. It includes Lewis' axioms, and results, but is broader because probabilities other than 1 are also accommodated; it corrects a misconception in Pearl's account, too. But above all, the whole of it is derived just by logic from standard definitions, and it shows thereby the intimate connection between probability theory and standard logic. Ellis looked forward to the day when the probability and predicate calculi would be unified into a single calculus, and we now come to see that the epsilon calculus enables us to do just that.