Radical Probabilism and Bayesian Conditioning*

Richard Bradley
Department of Philosophy, Logic and Scientific Method
London School of Economics
Houghton Street
London WC2A 2AE

December 5, 2005

Abstract

Richard Jeffrey espoused an anti-foundationalist variant of Bayesian thinking that he termed 'Radical Probabilism'. Radical Probabilism denies both the existence of an ideal, unbiased starting point for our attempts to learn about the world and the dogma of classical Bayesianism that the only justified change of belief is one based on the learning of certainties. Probabilistic judgement is basic and irreducible.

Bayesian conditioning is appropriate when interaction with the environment yields new certainty of belief in some proposition but leaves one's conditional beliefs untouched (the Rigidity condition). Although, Richard Jeffrey denied the general applicability of this condition, one of his main contributions to probabilistic thinking is a form of belief updating – now typically called Jeffrey conditioning or probability kinematics – that is appropriate in circumstances in which Rigidity is satisfied but where the interaction causes one to re-evaluate one's probability judgements over some partition of the possibility space without leaving one certain of the truth of any particular element. The most familiar occasion for Jeffrey conditioning is receipt of uncertain evidence: things partially perceived or remembered. But it also serves to illuminate belief updating occasioned by a change in one's degrees of conditional belief, a kind of belief change largely ignored by classical Bayesianism. I argue that such changes in conditional belief can also be basic (in the sense of not analysable as a consequence of conditioning on factual information) and offer a kinematical model for a particular kind change in conditional belief. Both are applied to changes in preference.

*This paper is dedicated to the memory of Richard Jeffrey, from whom I learnt so much.

1 Radical Probabilism

Richard Jeffrey's death two years ago deprived philosophy of a highly original thinker and exceptional human being. His writings span epistemology, the logic and philosophy of science, and both individual and social decision theory. In all of these fields he made important contributions, some widely recognised, some still waiting to be. Richard Jeffrey was a card carrying Bayesian, albeit of a somewhat heretical kind, both in epistemology and decision theory. His two most important ideas - the version of Bayesian decision theory that he developed with the help of Ethan Bolker and the theory of a probability kinematics, a generalisation of Bayesian conditioning to contexts in which the evidence is 'uncertain' - belong to these two faces of Bayesianism. It is the latter idea that is the focus of this essay, though connections to the former will be drawn. My aim is to situate the idea of a probability kinematics within Jeffrey's critique of Bayesianism and to demonstrate its fecundity by applying it to new problems, such as how to respond to evidence relevant to conditional belief and how to revise preferences.

Much of Jeffrey's work is shaped by his adherence to a variant of Bayesianism that he termed Radical Probabilism. Broadly a probabilist is someone who views doxastic judgement as a matter of adopting a probability assignment or, more generally, classes of assignments with particular characteristics. Jeffrey's own take on Probabilism is characterised by three features. It is, in the first place, thoroughly subjectivistic in its insistence that probability assignments represent *judgements*; things 'in the mind' not agent-independent features of the world.[1] Secondly, Jeffrey takes Probabilism to be a species of Pragmatism and, in particular, endorsed the view that degrees of beliefs should be understood in terms of their role as bases for, and determinants of, action.

> "I see the seventeenth century emergence of probability as a fission of the concept of approval into judgemental probability and judgemental desirability, with the second element coming into full view only lately, in the work of Ramsey ...". – [9, p. 66]

Versions of both of these views are commonplace in Bayesianism, of course, and subject to healthy debate. But it is the third 'anti-foundationalist' aspect of Jeffrey's thought that is of most concern here and where the departure from classical Bayesianism is most evident. For Bayesians, probabilistic judgement is a matter of drawing out the consequences of what one has learnt (the informational input or evidence) in the light of one's credal commitments (the prior belief state). For rationalistic Bayesians such as Laplace, Keynes, Carnap and Harsanyi prior belief states are themselves subject to norms of epistemic rationality, aimed for the most part at ridding the mind of bias of one kind or another. For them belief states are justified just in case they are achieved by valid inference from a privileged initial state and the totality of evidence yielded by experience. Subjective Bayesians, on the other hand, reject the idea of a rational prior belief state and so give up on the idea that there is a unique correct system of partial belief associated with any evidence base. The focus of attention switches to *changes* of belief and its drivers, analysed as one part input of information and one part inductive logic. For subjective Bayesians the fact that you believed X to degree a and now believe it to degree b is to be justified by the fact that what you have learnt together with your prior beliefs commits you to believing X to this degree.

Jeffrey's Probabilism is 'radical' in two ways. In the first place it rejects the idea that probabilistic judgement requires a bedrock of certainty and that belief change is justified only if required by the receipt of new information. Secondly he rejects the idea of an inductive logic and, in particular, the claim that updating by conditioning is obligatory on the pain of irrationality. These views are most concretely expressed in Jeffrey's development and interpretation of his probability kinematics, and it is to this that we now turn.

[1] In this regard Jeffrey was a devoted follower of De Finetti, both propounding and extending the latter's critique of the frequentist interpretation of probability and searching, somewhat less sucessfully, for a subjectivist reading of the objective chances postulated by Quantum Mechanics.

2 Classical Conditioning

For subjective Bayesianism the central problem of epistemology is how we should change our partial beliefs in response to new evidence, where evidence is to be thought of as information generated by interactions with the environment e.g. from observation, experimentation or the reliable testimony of others. Its formulation of the problem, and solution to it, is elegant and compelling. Suppose that your current degrees of belief are represented by probability measure p, defined on a Boolean algebra Ω of propositions representing all possibilities of concern to you, and that as a result of some such interaction with the environment you learn that A. In these circumstances your new degrees of belief, q, should go by your initial conditional degrees of belief given that A. i.e. $q(\cdot)$ should equal $p(\cdot|A)$.

Now, as Richard Jeffrey was fond of noting, adopting $p(\cdot|A)$ as your new degrees of belief is demonstrably the correct thing to do just in case, for all propositions $B \in \Omega$, both:

(i) *Certainty*: $q(A) = 1$

(ii) *Rigidity*: $q(B|A) = p(B|A)$.

So much is just a matter of probability theory. The important question is whether and when we can expect these two conditions be satisfied. The contention of classical Bayesianism in this regard is that they will be satisfied just in case A describes all and everything that is learnt by the agent as a result of the interaction with the environment. One qualification: it is commonly recognised that the Rigidity condition is liable to be violated when the manner in which the agent learns that A has some non-rational effect on the agent's attitudes; an effect that is, so to speak, independent of the propositional content of A. If, for instance, one learns of the consequences of excessive alcohol consumption by doing the drinking oneself or of the presence of a poisonous snake in the house by standing on it, there is every possibility that other beliefs (as well as one's non-credal attitudes) will be altered in the process and in a manner not representable as a conditioning on what has been learnt. So what the Bayesian model claims to describe is just the *rational* effects of learning something; what you are licensed by your current beliefs to infer from what you have learnt.[2]

This claim is typically defended by means of a dynamic Dutch Book argument, purporting to show that following the Bayesian rule of conditioning is a matter of pragmatic coherence. In the background, however, is the following intuitive argument. Suppose that A is all and everything that is learnt, that all changes to the agent's partial belief are rational effects of her learning that A, but that her new degrees of belief are not her old degrees of belief conditional on A. Then her conditional degrees of belief given A must have changed. But the truth of A is not itself a reason to change one's conditional beliefs given A, so something more than A must have been learnt. But that is contrary to the supposition that A is all that is learnt.

This argument is not entirely convincing as it stands: it is possible, for instance, that A *together* with other beliefs will justify a change to one's conditional beliefs given A. (More on this and the Dutch Book arguments later). But even if the Bayesian contention is true, it's scope is narrower than it might at first appear. People are often not aware of all that they have learnt or they fail to adequately represent it, and it is only the failure of the Rigidity condition that alerts us to this. This possibility is amply demonstrated by the Hangman or Monte Hall paradox.[3] More importantly the cases in which the Certainty condition is met do not exhaust motivated belief change.

Example 1 *I overhear a conversation in a foreign language and from the sounds of words and the mannerisms of the speakers I conclude that they are most likely, say, Spanish, but perhaps Catalan or even French. They seem to be assenting to each other's remarks by utterances of 'si', but perhaps I am mishearing. Could the 'si' be the French denial of a negated assertion? There is no hope here of*

[2] This is not the same as equating the partial beliefs formed by conditioning on what one has learnt with the beliefs one is justified in having. Sobering up may have a postive effect on one's beliefs (by bringing them closer to the truth) without this effect being a rational consequence of something that has been learnt.

[3] See Jeffrey [9, p. 122-4] for a discussion of this.

producing a sentence that summarises all and only the facts learnt (and believed with probability one) in the encounter. Relevant evidence is not entirely indubitable, many of the cues never make it into consciousness (and perhaps cannot) and I don't have well-defined conditional degrees of belief for the speaker's language, given the bits of evidence that do.

The possibility that someone's degree of belief in a particular proposition may change with reason without their being sure of either its truth or falsehood was first noted by Ramsey in his criticism of Keynes' interpretation of probability as a logical relation between propositions.

> "I think I perceive or remember something but am not sure; this would seem to give me some ground for believing it, ... He [Keynes] cannot justify a probable belief founded not on argument but on direct inspection. In our view ... there is no objection to such a possibility, with which Mr Keynes' method of justifying probable belief solely by relation to certain knowledge is quite unable to cope." [12, p.86]

Ramsey's observation is bad news not just for Keynes' theory, but for any epistemology that would have our probability judgements based on certainties. For cases of uncertain evidence, as we might call them, can be accommodated by the classical conditioning model only if one's new probabilities for the proposition about what one perceived or remembered are themselves products of conditioning on some 'deeper' evidence proposition of which one is certain. That such a proposition must always exist is what Jeffrey called the "empiricist myth of the sensuously given data proposition" [9, p. 3]. The myth has proven to be remarkably difficult to dispel, but myth it is. This is demonstrably so for cases of belief change in which we come to doubt propositions that we previously held to be certain: if A is the 'deep' proposition representing the sum total of what one has learnt, $p(B) = 1$ and $q(B) < 1$, then clearly $q \neq p(\cdot|A)$.[4] But our example suggests that it is true in general; that:

> "probabilistic judgement may be appropriate as a direct response to experience, underived from sure judgment that the experience is of such and such a character".[9, p. 45]

3 Probability Kinematics

In cases where one acquires 'uncertain evidence', Jeffrey offers a rule of conditioning that generalises the Bayesian one. Specifically suppose that $\{A_i\}$ is a partition, a set of mutually exclusive and exhaustive propositions, and that as a result of interaction with the environment (or indeed reflection or deliberation), the agent's probabilities for each A_i changes from $p(A_i)$ to $q(A_i)$. Then:

Definition 2 *An agent is said to obtain her new degrees of belief, q, by **Jeffrey conditioning** on the partition $\{A_i\}$ just in case q is related to p by the 'kinematical' formula:*

$$q(B) = \sum_i p(B|A_i).q(A_i)$$

An attractive feature of Jeffrey's kinematics is that it allows one to be a fallibilist about evidence and yet still make use of it. An apparent sighting of one's friend across the street, for instance, can be revised subsequently when you are told that he is out of the country. A closely related feature is the order-dependence of Jeffrey conditioning: conditioning on a particular redistribution of probability over a partition $\{A_i\}$ and then on a redistribution of probability over another partition $\{B_i\}$ will not in general yield the same posterior probability as conditioning first on the redistribution over $\{B_i\}$ and

[4] See Howson [8] for a full development of this point. A Bayesian might however take this as an argument against full belief in any contingent proposition.

then on that over $\{A_i\}$. This property, in contrast to the first, has been a matter of concern rather than admiration; a concern for the most part based on a confusion between the experience or evidence and its effect on the mind of the agent.[5]

Suppose, for instance, that I expect an essay from a student. I arrive at work to find an unnamed essay in my pigeonhole with familiar writing. I am 90% sure that it is from the student in question. But then I find that he left me a message the day before saying that he thinks that he may well not be able to bring me the essay in the next couple of days. In the light of all that I have learnt, I now lower to 30% my probability that the essay was from him. Suppose now I got the message before the essay. The final outcome should be the same, but I will get there a different way: perhaps by my probabilities for the essay coming from him initially going to 10% and then rising to 30% on finding the essay. The important thing is this reversal of the order of experience does not produce a reversal of the order of the probabilities: I do not think it 30% likely that I will get the essay after hearing the message and then revise it to 90% after checking my pigeonhole. *The same experiences have different effects on my probabilities depending on the order in which they occur.* (This is, of course, just a particular application of the rule that my posteriors depend both on the priors and the inputs). On the other hand, when we compare a situation in which the probability of X is revised to x_1 and then to x_2 with that in which it is revised to x_2 and then to x_1, we are generally dealing with cases where the experiences causing the probabilities to take these values must be different. So they will not be instances of order reversal of experience.[6]

What the order-dependence of Jeffrey conditioning (and the concern it causes) *does* reflect is the degree to which Jeffrey's radical probabilism departs from the empiricism of classical Bayesianism. Although formally Jeffrey conditioning is a generalisation of classical Bayesian updating, the epistemological doctrines associated with the latter are not so easily carried across. Classical conditioning is partially characterised by satisfaction of the Certainty constraint on the agent's posterior probabilities; a constraint which expresses the fact that some proposition A has been learnt i.e. that A is true and and fully believed to be so. This is captured by calling A the *input* into the agent's belief space. In the more general case the constraint on the agent's posterior degrees of belief takes the 'non-certain' form $q(A) = x < 1$. But although the constraint fixes the agent's new state of belief concerning A there is no implication that it is true that the probability of A is x and hence no implication that the agent has learnt that it is so. There is not even the implication that the revision brings the agent's beliefs closer to the truth about A: the evidence concerning A may accumulate, but its probability of truth need not.

All of this renders highly problematic the sense in which one can speak of *learning* by Jeffrey conditioning. For what is represented by the constraint on the agent's posterior degrees of belief is what the agent takes away from experience, her judgement upon it, rather than what experience delivers to her (with attendant connotations of objectivity). Whether the agent is justified or correct in adopting the constraint that forms the basis Jeffrey conditioning is a question that is not settled by the kinematical model itself. It is for this reason that the Reflection principle and related conditions of dynamic coherence are so unconvincing.[7] One may know that one's beliefs will change as a result of an experience without having any confidence that the change will be induced by something learnt. If I believe that I will come to learn that X then I have reason to believe X right now. But if I simply believe that I will come to believe that X then this gives may no reason to believe X (in the absence, at least, of the belief that my new belief will be justified).

Many Bayesians who accept the usefulness of Jeffrey's rule for cases of uncertain evidence find the abandonment of a notion of learning by conditioning unpalatable. It is not hard to understand why: it is the assurance that a policy of Bayesian conditioning brings one closer to the truth over time, backed by the famous convergence theorems, that makes its subjectivism consistent with confidence

[5] For a comprehensive discussion of these properties of Jeffrey conditioning, see Diaconis and Zabell [4].
[6] See Wagner [20] for further discussion.
[7] Such conditions have been postulated by, for instance, van Fraassen [19] and Skyrms [16].

that growth in our knowledge of the external world is possible.[8] One way of responding to the problem, what might be called the domestication strategy, is to try and identify an input from the world that externally justifies the constraint on the agent's new degrees of belief that serves as the basis for Jeffrey conditioning. Field [6], for instance, starts with the familiar Bayesian thought that Bayes or odds factors represent what is learnt from experience, independently of the agent's priors, and uses them to define an revision-prompting input parameter.[9] But Garber's [7] counterexample to his proposal shows that this does not work: the Bayes factors must be regarded, not as inputs, but as features of the agent's response to observation or experience. It does not follow from the fact that your Bayes factors represent what you have gleaned from observation, that they have the kind of objectivity which obligates others to modify their beliefs using them as constraints on their posteriors.[10]

The radical probabilist has little confidence in the domestication strategy and instead gives up on some of the foundationalist ambitions of Bayesians.

> Radical Probabilism makes no attempt to analyze judgement into a purely rational component and a purely empirical component, without residue. It rejects the ... picture of judgement as a coin with empirical obverse and rational reverse.[11]

The value of the kinematic model lies in the fact that it provides a framework in which probabilistic judgements can be made and policed. But whether the judgements so framed are any good will depend on the judgmental skills of the agent, typically acquired not in the inductive logic class but by subject specific training.

4 Revising Conditional Belief

Jeffrey conditioning on a partition $\{A_i\}$ is appropriate whenever redistribution of belief across $\{A_i\}$ leaves the agent's degrees of conditional belief unchanged i.e. when the Rigidity condition applies to all the conditional degrees of belief given A_i. This is just a matter of probability theory. What requires further consideration, however, is the extent to which it can be expected that Rigidity will apply: we turn to this in the next section. The contention that I want to consider here, however, is that updating on uncertain evidence does not exhaust all cases of motivated belief change lying outside of the scope of classical Bayesianism. In particular it is evident that not just our partial beliefs but our conditional beliefs too can change as a result of some kind of interaction with the environment. Here are some examples:

1. *Observation*: Testing for the sex ratios of, say, pike in some particular lake by means of random sampling of fish can lead to revisions to one's probability for a fish being a male, given that it is a pike.

2. *Experimentation*: I am having trouble opening a lock on the front door, but know I have the right key. After much fiddling, I conjecture that if I pull the door towards me while turning the key, the lock will open. After repeated trials, I am sure that I am basically right, even though it does not work on every attempt. My conditional probabilities for someone opening the door given that they follow this procedure rise near to one.

---

[8] See Savage [13, pp 46-49] and Earman [5] for respectively a presentation and critical discussion of these convergence results.

[9] The Bayes factor for propositions A and B is the ratio of new odds to old odds: $\frac{q(A)}{q(B)}/\frac{p(A)}{p(B)}$.

[10] Though sometimes it may be reasonable to do so; namely when I trust your response to experience, but do not share your priors. (See Jeffrey, [9, pp. 7-9]).

[11] Jeffrey [9, p. 3]

3. *Testimony*: I consult an oracle to find out whether I will succeed at my driving test tomorrow. The oracle tells me that if the examiner has a moustache then I will pass. Since I trust the oracle my conditional probability for passing given that the examiner has a moustache goes to one.

Classical Bayesians must represent such changes in conditional probabilities as consequences of learning the truth of some set of evidence propositions. This is plausible in some cases. The belief change induced by the sampling of fish to test for sex-ratios of pike, for instance, might be the result of conditioning on a data proposition believed to degree one concerning the proportion of pike in the sample that were found to be male. The other cases are less amenable to treatment of this kind, however. Although I believe the conclusion I reach about how to open the door is right in virtue of certain facts about the position of the door, the state of the lock and so on, I do not know what these facts are. There are also some facts of which I am aware, in particular those relating to my actions and their outcome, that no doubt played a role in my coming to believe that the lock will open if I pull the door while turning the key. But though my judgement is made in light of these facts, it is not necessitated by them in any apparent way.

In the third case it might seem natural to represent my belief change as an instance of conditioning on a proposition gleaned from the oracle's testimony. The problem is identifying the right proposition to condition on. Although what the oracle tells me implies that either the examiner will not have a moustache or I will pass, my new probabilities need not equal my prior conditional probabilities given the truth of this proposition. I might, for instance, judge that the oracle's pronouncement gives me no reason to change my beliefs about the likelihood of a moustached examiner (I form my beliefs about this from the statistics say). Then although my probability for getting a moustached examiner and passing is brought into line with my probability for moustached examiners, the latter remains unchanged. Perhaps the proposition we have been considering is not strong enough. Certainly I have learnt other things: that the oracle has spoken, that I have heard her, etc. but it is hard to see what the relevance of these to my probabilities for moustached examiners should be.

Our testimony example belongs to an particularly interesting set of cases in which the interaction with the environment gives us cause us to change one or more of our conditional beliefs, given some possibility A, without it giving us cause to change our probabilities for A. Changes of belief taking this form I will call updating by Adams conditioning, because Ernst Adams' theory of conditionals suggests that this is an appropriate way to respond to the truth of a conditional sentence.

Definition 3 *Let A and B be propositions such that $1 > p(B|A) > 0$ and suppose that the agent is caused to change her conditional degrees of belief for B given A from $p(B|A)$ to $q(B|A)$. Then her new partial beliefs, q, are said to be obtained from p by **Adams conditioning** on this change in conditional probabilities just in case:*

$$q(X) = p(ABX).\frac{q(B|A)}{p(B|A)} + p(A\neg BX).\frac{q(\neg B|A)}{p(\neg B|A)} + p(\neg AX)$$

Example 4 *I have arranged to meet a friend at a cafe and I am rushing to meet him. Then I hear on the radio that there is a jam on the road he would normally take to get there and that cars are not moving at all. My conditional belief for (B) my friend arriving late, given that (A) he takes his usual route, now goes to one. If I Adams conditioning on my new conditional beliefs, then for all prospect X, $q(X) = p(X|AB).p(A) + p(X|\neg A).p(\neg A)$. And in particular: $q(B) = p(A) + p(\neg AB)$. As $p(A)$ is high, I decide not to hurry.*

What makes Adams conditioning salient is that, in a certain sense, it is the exact compliment of Jeffrey conditioning. For note that it follows immediately from its definition that $q(\neg A) = p(\neg A)$ and hence $q(A) = p(A)$. So in Adams conditioning it is the conditional probabilities with respect to elements of a partition that change while the probabilities of the elements themselves remain rigid, rather than the other way round. Consequently study of this kind of revision offers the possibility of extending kinematical modelling to cases where interaction with the environment affects both the agent's conditional beliefs and her conditional ones, by representing them in terms of combinations of Jeffrey and Adams conditioning.

In another sense, Adams conditioning appears to be just a special case of Jeffrey conditioning. Suppose for instance that your conditional probability for B given A changes, but your probability of A does not. Your new belief state is then just the one you would obtain by Jeffrey conditioning on the partition $\{AB, A\neg B, \neg A\}$ in case the sum of your new probabilities for AB and $A\neg B$ equals your old probability for A i.e. $q(AB) = q(B|A).p(A)$, $q(A\neg B) = q(\neg B|A).p(A)$ and $q(\neg A) = p(\neg A)$. This follows immediately from Theorem 5 below (proof in the appendix).

Theorem 5 *An agent's new degrees of belief, q, are obtained from her old, p, by Adams conditioning on a change in the conditional probability of B given A iff for all propositions X:*
(i) Independence: $q(A) = p(A)$
(ii) Rigidity: $q(X|AB) = p(X|AB)$, $q(X|A\neg B) = p(X|A\neg B)$, $q(X|\neg A) = p(X|\neg A)$

Let us be clear about the import of this result however. The point is not that changes in conditional beliefs are after all a kind of updating on uncertain evidence. It is rather that the kinematical model is flexible enough to represent very different kinds of effects of interaction with the environment, including the acquisition of uncertain evidence and learning conditional relationships between possibilities, as well as the outcomes of revision-inducing processes not discussed here (such as reflection or deliberation). When we talk of Jeffrey or Adams conditioning we implicitly address ourselves to a kinematical model of belief change with a particular kind of motivation or source for the change.

Theorem 5 in fact establishes a specific case of an general truth about the kinematical model; namely that in marked contrast to classical conditioning *any* revision of probabilities defined on a countable set of propositions can be represented as an instance of Jeffrey conditioning on some partition satisfying the Rigidity condition (see van Fraassen [18] and Diaconis and Zabell [4] for further discussion of this feature). Important though this result may be, however, the fact that revisions can be represented in a certain kind of way does not imply either that this actually was the way the agent revised her beliefs or, more importantly, that this is the way that she should have. A representation result of this kind cannot settle the question of which partition of possibility space is the one with respect to which conditional degrees of belief should remain invariant (and so, for instance, whether or when Adams conditioning is the appropriate response to a change in conditional belief). In the next section, we address the question directly.

5 Revising Preferences

I now want to consider the question of how we should revise our preferences in the light of experience. Jeffrey's own discussion of preference change is largely confined to a small paper "*The Kinematics of Preference*" penned in response to a problem set for him by Wolfgang Spohn and Ethan Bolker.[12] But many of the tools we will require for our investigation are present in his work. Broadly speaking changes in preference have two sorts of causes: changes in beliefs and what might be called changes in tastes. I shall confine attention to the former and in particular, since it leads on most naturally from what we have discussed thus far, to cases where preference change is induced by a redistribution of partial belief across some particular partition of the possibility space.

To do this a bit more formally, I will make use of the decision theory developed by Richard Jeffrey in his '*Logic of Decision*' [10]. In his framework the state of mind of a (maximally opinionated) rational agent is represented by a pair of functions, $< p, v >$, defined on a Boolean algebra of propositions,

[12] Reproduced in [9].

Ω, and such that p is a probability measure of her degrees of belief and v a real-valued (desirability) measure of her degrees of preference, satisfying the following condition. Let $\{A_i\}$ be any partition of Ω. Then if $p(X) \neq 0$:

$$v(X) = \sum_i v(XA_i).p(A_i|X)$$

Suppose that an agent's initial state of mind is represented by the pair $< p, v >$. Suppose that as a result of interaction with the environment the agent's degrees of belief for the elements of some partition $\{A_i\}$ change, and that after Jeffrey conditioning on these changes her new degrees of belief are represented by the probability function, q. How should the agent's preferences change as a result? To isolate the impact on her preferences of the changes in her partial belief, we will suppose that the interaction has no independent effect on her preferences other than those resulting from her changes in belief. Then I suggest that her new preferences for any prospect X can be obtained from her old preferences for the XA_i by combining them with her new conditional probabilities for the A_i given X.

Since desirability representations of preferences are not unique, there will be more than one way of expressing this thought in terms of a numerical relation between an agent's old and new desirabilities following a change in her partial belief. I do think there is a 'best' expression, but the argument for it will take us too far afield. It would be more useful to demonstrate how this kind of preference change works by means of a couple of examples. The first illustrates the effect on the desirability of A of a change in the conditional probability, given A, of a prospect that matters to the agent.

Example 6 *Suppose that I prefer drinking white wine to red. But then I read in the papers that drinking red wine (R), but not white (W), reduces the chances of a heart attack (H). To form my new preferences between drinking red or white wine I recalculate the desirability of these prospects using my old degrees of desire, v, for RH, R¬H, WH and W¬H and my new conditional probabilities for a heart attack given my consumption of either kind of wine. Then I should prefer red to white just in case:*

$$v(RH).q(H|R) + v(R\neg H).q(\neg H|R) > v(WH).q(H|W) + v(W\neg H).q(\neg H|W)$$

If the probability of a heart attack is sufficiently lessened by drinking red wine as opposed to white, my preferences between the two may well switch.

The second kind of belief change relevant to preference is when some prospect A becomes more or less attractive as result in a change in the probability of some other possibility B, not because of any probabilistic dependence between the two, but because of the attractiveness of A depends on whether B is the case or not. To isolate this effect we consider an example in which the relevant prospect is probabilistically independent of the possibilities that condition the desirability of its consequences.

Example 7 *Suppose that, in the expectation of a sunny day, I am planning (B) a trip to the beach. However, after listening to the weather report, I change my mind: it now seems much more likely that it will be (R) a rainy day than (S) a sunny day. To form my new preferences I take my old degrees of desire for the prospects SB, RB, S¬B and R¬B and weight them by my new probabilities for a sunny or rainy day. Then I should take the trip if:*

$$v(SB).q(S) + w(RB).q(R) > v(S\neg B).q(S) + w(R\neg B).q(R)$$

Assuming that SB is preferred to S¬B and R¬B to RB then I should abandon the trip if my new probabilities for rain are too high.

In both examples we have implicitly assumed that a shift in the agent's degrees of beliefs for the elements of some partition $\{A_i\}$ does not affect her relative degrees of desire for the XA_i. This assumption is in fact stronger than is required for preference change to be properly characterised as Bayesian.

What matters is that the agent's conditional preferences for prospects, given the A_i, should not change. Conditional preferences are judgements of preference made under the supposition that some condition is true. Thus an agent has a conditional preference for some prospect X over another prospect Y, given that A, when, having supposed that A is the case, she finds that she would prefer that X to that Y. When this judgement is unaffected by a change in her attitude to A, then we say that her conditional preference for X over Y is *rigid* with respect to A.

Although not always explicitly acknowledged, assumptions about the rigidity of conditional preferences (in particular conditional preferences for money) play an important role in the dynamic Dutch book arguments standardly employed in justifying Bayesian conditioning and which are discussed in the next section. Here we show that rigidity of conditional preference is both necessary and, jointly with the Bolker-Jeffrey axioms of preference, sufficient for the validity of Jeffrey conditioning. We will do so by making use of Bolker representation theorem for Jeffrey's decision theory, but to avoid needless complication associated with the lack of uniqueness of probability representations of preference in the Jeffrey-Bolker framework we will assume that the agent's preferences admit only of unbounded desirability representations (we say her preferences are unbounded in this case). In this case it follows from Bolker's theorem that subjective probabilities are uniquely determined by rational preference and desirabilities are determined up to a choice of scale.[13]

Let Ω be a Boolean algebra of prospects with the contradictory proposition F removed. Let $\geq$ and $\geq^*$ respectively represent the agent's preferences before and after her changes in belief regarding the A_i. Let $\geq_{A_i}$ and $\geq^*_{A_i}$ be her corresponding conditional preferences on the supposition that A_i is true. Then we postulate:

Axiom 8 *(Rigidity of Conditional Preference)* $\forall X, Y : XA_i \neq F,\ YA_i \neq F,\ X \geq_{A_i} Y \Leftrightarrow X \geq^*_{A_i} Y$

Theorem 9 *(Representation of Jeffrey Conditioning) Assume that* $\geq$ *and* $\geq^*$ *are unbounded and satisfy both the Jeffrey-Bolker axioms of preference and the axiom of Rigidity of Conditional Preference. Let* $< p, v >$ *and* $< q, w >$ *be pairs of probability and desirability functions that respectively represent* $\geq$ *and* $\geq^*$*. Then* q *is obtained from* p *by Jeffrey conditioning on the partition* $\{A_i\}$*.*

We have already discussed some of the contexts in which the rigidity condition for conditional belief can fail, including cases where A's turning out to be true has some 'non-rational' effect on her attitudes. We can expect that these will be cases in which the rigidity condition on conditional preference will fail as well, an expectation borne out by the earlier examples of learning about the effects of alcohol by drinking and of the presence of a snake by standing on it. In any case, I do not intend that the condition be thought of as rationally binding on agents. Nor should the representation theorem for Jeffrey conditioning be thought of a justifying in any kind of general way this form of belief revision. What it does do is give a different kind of answer to the question of when is it correct to update one's degrees of belief by Jeffrey conditioning on some partition: namely, when experience motivates a change in one's attitudes to the elements of the partition, but leaves one's conditional preferences given any one of them unchanged. Note, finally, that rigidity of conditional preference is not necessary for Jeffrey conditioning. But when conditional preferences given the elements of the partition $\{A_i\}$ change, following some redistribution of probability over the A_i, then either this redistribution does not adequately represent the informational effects of experience or some of the changes in the agent's preferences have a non-informational origin e.g. from a change in taste. Such changes are non-rational rather than irrational.

[13] See Bolker [2].

6 Rigidity and Dynamic Coherence

Earlier on an intuitive argument for classical Bayesian conditioning was presented that rested on the claim that coming to believe a proposition does not in itself give one reason to change one conditional degrees of belief given its truth. An analogous argument could be given for both Jeffrey and Adams conditioning. If what one learns from an interaction with the environment is appropriately represented by a shift in one's degrees of belief for some possibility A, or by a shift in one's conditional degrees of belief for some B given A, and by nothing more and nothing less than these shifts, then the interaction has furnished no reason for a change in one's conditional degrees of belief given A or, in the other case, of one's unconditional degrees of belief for A. Else the effect of the interaction was not properly represented. More than this is needed however to support the claim that Jeffrey and Adams conditioning are universally valid forms of revision. It must also be the case that the primary change in one's probabilities (or conditional probabilities) should not serve as a basis for an inference that leads to revision of one's conditional probabilities (or unconditional ones). To put it differently, the argument for both forms of conditioning presupposes that an agent's degrees of belief in some A and her conditional degrees of belief given A are epistemically independent of one another in the sense that a change in one does not in itself motivate a change in the other, *irrespective of what else the agent believes.* I doubt that this is generally the case.

The informal argument for conditioning presented above leaves the notion of 'furnishing a reason for belief change' much too vague, and Bayesians more usually depend on dynamic Dutch Book arguments. These arguments show that an agent who commits herself to any policy for revising her beliefs other than Bayesian conditioning is vulnerable to a sure loss from acceptance of a finite set of bets that she considers fair (relative to her degrees of belief). It follows that if she commits herself to credal changes that amount to anything more or anything less than that she has reason to (in the informal sense), she will find that she has pragmatic grounds for regretting her commitments.

Recall how this argument goes for classical conditioning.[14] Suppose that the agent has degrees of belief p and that she commits herself to a revision policy such that for some A and B, $p(B|A) - q(B) > d > 0$, where q represents her degrees of belief after learning that A. Then a bookie can make a sure gain by selling:

(i) a fair conditional bet on B in the event of A costing \$$p(B|A)$ and that pays \$1 if AB, nothing if $A\neg B$, and is called off (with the stake returned) if $\neg A$;

(ii) a fair bet on A costing \$$d.p(A)$ that pays \$$d$ if A and nothing otherwise; and in the event of A,

(iii) a fair bet on $\neg B$ costing $q(\neg B)$ and paying \$1 if $\neg B$ and nothing otherwise.

Then in the event of A the agent loses \$$(p(B|A) - q(B) - d) > 0$ and in the event of $\neg A$, she loses \$$d.p(A)$.

A similar defence can be given for the claim that if the effect of interaction with the environment is appropriately represented by a redistribution of the agent's degrees of belief over a partition $\{A_i\}$ then, on pain of the sort of incoherence that renders one vulnerable to a sure loss, the agent should update their degrees of belief by Jeffrey conditioning on the partition $\{A_i\}$. A number of different versions of the argument exist, but the simplest one goes as follows.[15] Suppose that an agent's degrees of beliefs are represented at time t_0 by probability p and that at t_1 she undergoes some interaction with the environment that yields information about a partition $\{A_i\}$ and only about this partition. Soon thereafter, at time t_2, she learns the true member of the partition. Suppose that in the interval

[14] The original dynamic Dutch book argument for classical conditioning is due to David Lewis, reported in Teller [17]. For criticism of these arguments see Earman [5] and Howson [8].

[15] See, for instance, Armendt [1] and Skyrms [14]. The argument given here essentially follows that given by Skyrms [16].

between t_0 and t_2 she learns nothing about propositions in any other partition. Let q represent her degrees of belief at t_1 and r her degrees of belief at t_2. Now by the dynamic Dutch book argument for classical conditioning, if the agent is dynamically coherent then $r(\cdot|A_i) = p(\cdot|A_i)$ and $r(\cdot|A_i) = q(\cdot|A_i)$. So $q(\cdot|A_i) = p(\cdot|A_i)$ and, hence, q is obtained from p by Jeffrey conditioning.

What are we to make of these arguments? The dynamic Dutch Book argument for classical conditioning shows that under the assumed conditions - commitment to a belief revision policy which one is prepared to make public, a degree of self-consciousness about the changes in one's beliefs and a readiness to accept any fair bet - failure to update by conditioning leaves one vulnerable to a sure loss. But even when these conditions hold, it does not follow that classical conditioning is the uniquely rational updating strategy. Indeed, in cases where the interaction with the environment produces uncertain evidence regarding other propositions, conditioning on the strongest proposition learnt in the interaction will lead the agent into *synchronic* inconsistency. For instance, suppose that A is the strongest proposition learnt and that this learning is accompanied by a motivated re-evaluation of the probability of the elements of the partition $\{B, \neg B\}$ without the probability of either going to one e.g. when having an event (A) recounted to her vaguely reminds the agent of something that she saw or heard in the past (B). Then unless her new degree of belief for B happens to equal $p(B|A)$, she cannot, on pain of inconsistency, condition on A. Of course, if she doesn't she will be vulnerable to a Dutch Book. But this vulnerability is not a symptom of 'dynamic incoherence', at least if by this is meant 'irrationality', but a reflection of the gross imbalance of power between her and the bookie: the bookie can pick and choose which bets to make in accordance with what she knows about the agent's beliefs and her policies for changing them. A rational agent should refuse the bookie's game and not declare a revision policy. Indeed, to have such a policy seems very unwise, given that how one should respond to the information that A depends of what other shifts in one's degrees of beliefs are motivated by the process of learning that A.

For the Bayesian who accepts the possibility of motivated belief change not rooted in the receipt of certain information the natural response is to treat the argument for classical conditioning as a special case of the Dutch Book argument for Jeffrey conditioning. The latter takes as its premise that all motivated changes in belief are appropriately represented by a redistribution of probability across some partition of the possibility space. The problem now is that there is a way of reading this premise which makes it so strong that it cannot fail to follow that Jeffrey conditioning is the right way to update. For suppose that the effect of interaction on the agent's beliefs is appropriately represented by a shift in probability over the partition $\{A, \neg A\}$. Suppose also that she does not update by Jeffrey conditioning. It follows that her conditional degrees of belief given A must have changed as a result of the interaction and, hence, contrary to assumption, there has been some shift in probability over the partition $\{AB, A\neg B, \neg A\}$ not represented by the shift over $\{A, \neg A\}$.

This is surely not what the Bayesian has in mind: what we want to say is that if what the agent learns from the interaction is restricted to a certain partition, or that if "she believes with probability one that the learning experience only gives information about the partition in question"[16] then her conditional beliefs given the elements of that partition should not change. This requires that we can clearly separate the primary effect of experience (on some restricted partition) from the propagation of its implications through conditioning. But this is not without its difficulties for, as we have already seen, it is not possible to reduce the primary effect of experience to an input from the environment. The constraint on the agent's posterior degrees of belief that serves as the basis of Jeffrey conditioning is a *judgement*, informed by experience but also by the agent's beliefs. But then it follows from the fact that the agent's judgements are interconnected that the relevance of experience cannot be isolated to a single partition and that the agent should not believe it to be so.

The best we can hope to do, I think, is to draw a rough distinction between the changes directly motivated by what the agent observes or experiences and those arrived at by inference from them (recog-

[16] Skyrms [16, p. 288]

nising that the distinction is far from absolute). The independence claim will now read as postulating that if the direct effect of experience can be isolated to a change in the agent's probabilities over some partition, then whatever changes that are produced by inference from the primary effect should not include changes to the agent's conditional probabilities given the elements of the partition. And vice versa.

Unfortunately this claim is false: there are cases where an agent's partial beliefs over some partition and her conditional beliefs with respect to it are not independent of one another in this sense. The most common examples perhaps are those involving actions, when learning that the action has certain effects makes its performance more or less probable and when learning that it is likely to be performed raises the conditional probability of it having certain effects. An increase, for instance, in the degree to which I believe that I will be late for an appointment, given that I take the bus, is liable to make it less likely that I will in fact take it and thereby to motivate a decrease in the degree to which I believe that I will take it. Likewise, if I gather that someone is going to take the bus, then I might infer that he has knowledge about the reliability of the bus that makes the probability of him being late for an appointment, given that he takes the bus, rather low (lower, anyway, than I previously believed).

Examples involving actions are common, but they by no means exhaust the cases that exhibit the interdependence of conditional and unconditional belief.

Example 10 *Suppose that grain is stored in a chamber with two outflow pipes, A and B (see figure 1). Pipe A empties into a second chamber, X, while pipe B empties into both chamber X and another chamber, Y. Suppose that we know the amount of grain stored in the first chamber as well as the amount that ends up in chamber X when the outflow pipes are opened so that we are able to estimate with accuracy the probability of any piece of stored grain ending up in chamber X. Say its one half. We also know that any grain that goes down A ends up at chamber X, but we can only roughly estimate from the size and position of the pipes the probability that it goes down pipe A or B, and with what probability it will end up in chamber X in the event of it going down pipe B. Suppose now that sampling from the mouth of pipe B shows that we need to revise the degree to which we believe that a stored piece of grain will go down this pipe from, say, one third to one quarter. If our conditional degrees of belief in the grain landing up in chamber X given that it goes down B stays the same then as a matter of consistency we should change our probability for the grain ending up in chamber X. But this would be the wrong thing to do, since we already know how much grain lands up there. So our conditional degrees of belief ought to change in response to the change in probability for the grain going down pipe B, contrary to the conditioning model.*

We could treat the situation described in Example 10 as a change in the probability that the grain goes down pipe B accompanied by a change in the conditional probability that it will land up in chamber X given that it goes down pipe B. This would make it amenable to the kind of extended kinematical modelling fleshed out above. But to do this would be to ignore the fact that the change in conditional belief occurs *because* of the change in belief and does not merely accompany it. Similarly we could imagine a variation on the example in which we learn something about the likelihood of grain landing up in chamber X, given that it goes down pipe B, and then updating our probabilities for it going down pipe A as a result. This variation would be indistinguishable kinematically from the first case, but intuitively is quite different.

At face value this example refutes the Bayesian's claim that a change in someone's degrees of partial belief for some prospect does not give them reason to change their conditional degrees of belief given the truth of the prospect. Faced with a revised probability for the grain going down pipe B, we can opt to revise our degree of belief for the grain landing in chamber X *or* our conditional degree of belief in landing there given that it goes down pipe B *or both.* In general how we choose to accommodate a change in our partial belief in prospect A depends, for each prospect B, on how entrenched (to use a term from AGM belief revision theory) our degrees of belief for B are relative to our conditional degrees of belief for B

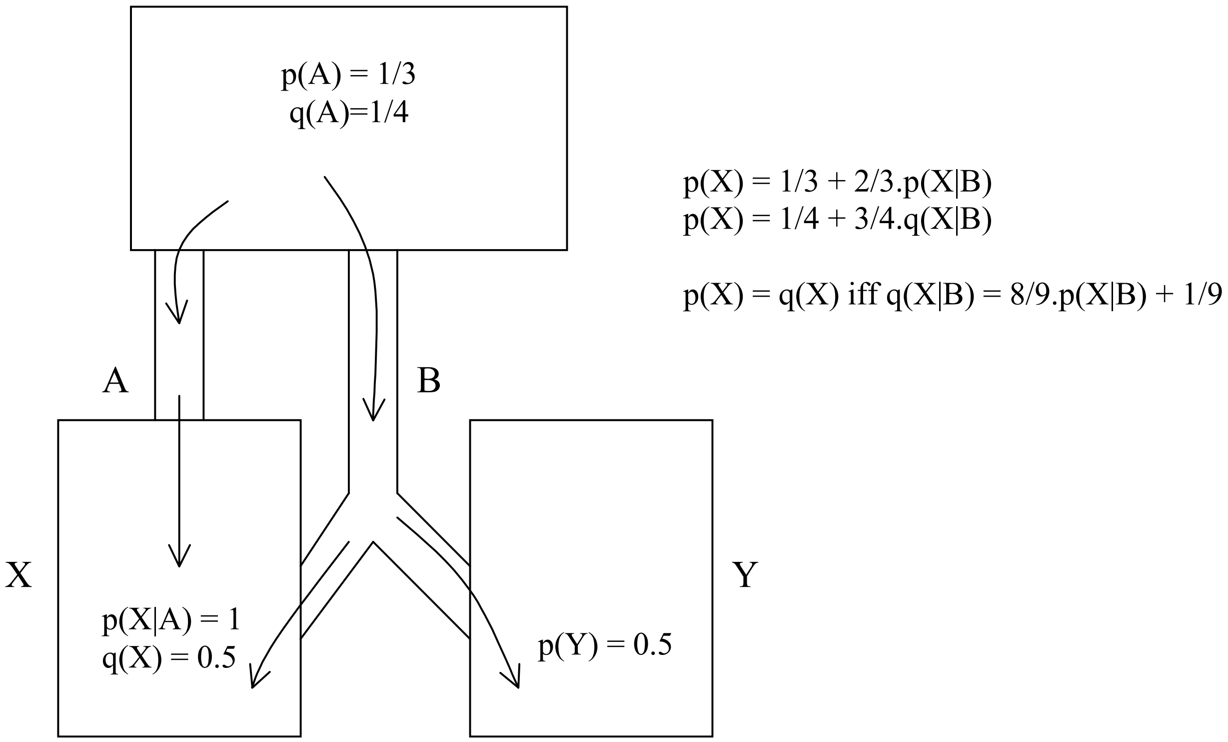
p(A) = 1/3
q(A)=1/4
p(X) = 1/3 + 2/3.p(X|B)
p(X) = 1/4 + 3/4.q(X|B)
p(X) = q(X) iff q(X|B) = 8/9.p(X|B) + 1/9
A
B
X
Y
p(X|A) = 1
q(X) = 0.5
p(Y) = 0.5

Figure 1:

given A. Probability theory alone cannot determine how to make the accommodation: some judgement must be made as to the epistemic standing of both our conditional and our unconditional beliefs. This is not to dismiss Bayesian conditioning, but to emphasise that it should not be thought of as a universal and mechanical rule of updating, but as a technique to be applied in the right circumstances, as a tool in what Jeffrey terms the 'art of judgement'.

7 Appendix: Proofs

Proof of Theorem 5. If q is obtained from p by Adams conditioning on a change in the conditional probability of B given A then by definition:

$$\begin{aligned} q(A) &= p(AB).\frac{q(B|A)}{p(B|A)} + p(A\neg B).\frac{q(\neg B|A)}{p(\neg B|A)} \\ &= p(A).q(B|A) + p(A).q(\neg B|A) \\ &= p(A) \end{aligned}$$

and:

$$\begin{aligned} q(X|AB) &= \frac{q(XAB)}{q(AB)} \\ &= \frac{p(XAB).\frac{q(B|A)}{p(B|A)}}{p(AB).\frac{q(B|A)}{p(B|A)}} \\ &= p(X|AB) \end{aligned}$$

Similarly $q(X|A\neg B) = p(X|A\neg B)$, $q(X|A) = p(X|A)$. So both the Independence and Rigidity conditions hold. Now suppose that the conditions hold. Then:

$$\begin{aligned} q(X) &= q(X|AB).q(AB) + q(X|A\neg B).q(A\neg B) + q(X|\neg A).q(\neg A) \\ &= p(X|AB).q(AB) + p(X|A\neg B).q(A\neg B) + p(X|\neg A).p(\neg A) \\ &= \frac{p(XAB)}{p(AB)}.q(AB) + \frac{p(XA\neg B)}{p(A\neg B)}.q(A\neg B) + p(X\neg A) \\ &= p(XAB).\frac{q(B|A)}{p(B|A)} + p(XA\neg B).\frac{q(\neg B|A)}{p(\neg B|A)} + p(X\neg A) \end{aligned}$$

■

Proof of Theorem 9. Let A be any member of $\{A_i\}$. Let $v(\cdot|A) =_{def} v(\cdot A) - v(A) + v(T)$. By Theorem 5 of Bradley [3, p.36], $< p(\cdot|A), v(\cdot|A) >$ represents $\geq_A$ over Ω. But by the rigidity of conditional preference, $\geq_A \equiv \geq^*_A$. So $< p(\cdot|A), v(\cdot|A) >$ represents $\geq^*_A$ over Ω. Then in view of the assumption that $\geq$ and $\geq^*$ are unbounded it follows from Theorem 7 of Bradley [3, p. 38] that there exists a pair of probability and desirability measures, $< q, w >$, that represents $\geq^*$over Ω and such that $w(\cdot|A) = v(\cdot|A)$, where $w(\cdot|A) =_{def} w(\cdot A) - w(A) + w(T)$. Finally it follows from Theorem 2 of Bradley [3, p. 33] that $q(\cdot|A) = p(\cdot|A)$. ■

References

[1] Armendt, B. (1980) "Is there a Dutch Book Theorem for Probability Kinematics?", *Philosophy of Science* 47: 563-588

[2] Bolker, E. (1966) "Functions Resembling Quotients of Measures", *Transactions of the American Mathematical Society* 124: 292-312

[3] Bradley, R. (1999) "Conditional Desirability", *Theory and Decision* 47: 23-55, Kluwer Academic Press

[4] Diaconis, P. and Zabell, S. (1982) "Updating Subjective Probability", *Journal of the American Statistical Association* 77: 822-30

[5] Earman, J. (1992) *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*, MIT Press

[6] Field, H. (1978) "A Note on Jeffrey Conditionalization", *Philosophy of Science* 45: 361-67

[7] Garber, G (1980) "Field and Jeffrey Conditionalization", *Philosophy of Science* 47: 142-5

[8] Howson, C. (1996) "Bayesian Rules of Updating", *Erkenntnis* 45: 195-208

[9] Jeffrey, R. (1992) *Probability and the Art of Judgement*, Cambridge University Press

[10] Jeffrey, R. C. (1983) *The Logic of Decision*, 2nd ed, Chicago, University of Chicago Press

[11] Joyce, J. (1999) *The Foundations of Causal Decision Theory*, Cambridge University Press

[12] Ramsey, F. P. (1926) "Truth and Probability" in D. H. Mellor (ed) *Philosophical Papers*, Cambridge: Cambridge University Press, 1990.

[13] Savage, L. J. (1972). *The Foundations of Statistics*, 2nd ed, Dover, New York

[14] Skyrms, B. (1987) "Dynamic Coherence and Probability Kinematics", *Philosophy of Science* 54: 1-20

[15] Skyrms, B. (1990) *The Dynamics of Rational Deliberation*, Harvard University Press

[16] Skyrms, B. (1996) "The Structure of Radical Probabilism", *Erkenntnis* 45: 285-297

[17] Teller, P. (1973) "Conditionalization and Observation", *Synthese* 26: 218-258

[18] van Fraassen, B. (1980) "Rational Belief and Probability Kinematics", *Philosophy of Science* 47: 165-87

[19] van Fraassen, B. (1984) "Belief and the Will", *Journal of Philosophy* 81: 235-256

[20] Wagner, C. (2002) "Probability Kinematics and Commutativity", *Philosophy of Science* 69: 266-278