

# Revising Incomplete Attitudes

Richard Bradley  
Department of Philosophy, Logic and Scientific Method  
London School of Economics  
Houghton Street  
London WC2A 2AE  
r.bradley@lse.ac.uk

May 6, 2009

## Abstract

Bayesian models typically assume that agents are rational, logically omniscient and opinionated. The last of these has little descriptive or normative appeal, however, and limits our ability to describe how agents make up their minds (as opposed to changing them) or how they can suspend or withdraw their opinions. To address these limitations this paper represents the attitudinal states of non-opinionated agents by sets of (permissible) probability and desirability functions. Several basic ways in which such states of mind can be changed are then characterised and compared with those found in AGM style models of attitude revision. Finally these models are employed to describe how agents make up their mind when deliberating.

## 1 Introduction

The agents populating Bayesian decision theoretic models are typically rational, logically omniscient and maximally opinionated. Rational in that their attitudes - beliefs, desires and preferences - are consistent both in themselves and with respect to one another; logically omniscient because they believe all logical truths and endorse all the logical consequences of their attitudes; and opinionated because they have determinate belief, desire and preference attitudes to all prospects under considerations. For many purposes these assumptions serve to simplify analysis without distorting it in any fundamental way. But most Bayesians would recognise that this package of assumptions can be criticised on both empirical and normative grounds, without necessarily agreeing on how severe these limitations are or on how they should be addressed.

In this paper, I want to focus on the assumption that agents are maximally opinionated which, in my opinion, is the most problematic of the three. There is no doubt that much has been learnt by studying opinionated agents - the

use of Dutch Book arguments to derive the additivity of rational degrees of belief is a case in point. But real people often do not, cannot, or will not reach judgements about some things. Furthermore there is little plausibility to the claim that rationality requires them to do so. Although the assumption that weak preferences are complete (i.e. that for every two prospects, one is more preferred than the other or they are regarded indifferently) plays an important role in the foundational representation theorems of decision theory, its role is that of making it possible to explore the consequences of rationality assumptions, such as transitivity and independence, in a simplified setting and not to express a further condition of rationality.<sup>1</sup>

The assumption that agents are opinionated about everything (the completeness assumption) also limits our ability to examine some of the different ways in which the attitudinal state of an agent can change. For one thing, it leaves no space open for the analysis of cases in which agents make up their minds, rather than simply change them, because belief and preference formation is in effect reduced to belief and preference revision. It also leaves no room for the treatment of cases in which an agent has reason to suspend judgement on some question or to retreat from a previously held attitude. This can happen for instance when someone discovers that the evidence supporting one of her beliefs is unreliable (prompting retreat) or when two experts are found to disagree on some question (prompting suspension). Finally, combined with the assumption that the agent is logically omniscient, it leaves no space open for the treatment of deliberation, understood as a process by which an agent derives the view that she should adopt with regard to some prospect in her attention from her attitudes to other related prospects.

These limitations are not common to all theories of attitude change. For instance, in the AGM approach to belief revision pioneered by Alchourrón, Gärdenfors and Makinson, belief states are represented by sets of sentences satisfying conditions of logical consistency and closure under deduction but explicitly *not* required to contain every sentence or its negation (see [6]). This allows for a characterisation not only of belief revision, but also of what is termed belief expansion and contraction; operations applicable respectively to cases in which an agent acquires a belief regarding some matter about which she had no prior attitude and in which the agent has some reason for abandoning or withdrawing her prior judgements. This is one reason to favour this approach over the Bayesian one. On the other hand, to my mind, AGM models are handicapped by the paucity of information that they contain about the strength of agents' attitudes, a feature which hinders their ability to connect to folk-psychological and social scientific explanations of behaviour that exploit these strength characteristics. It is difficult to see how such limitations in the AGM approach could be overcome without adopting something sufficiently like the Bayesian framework

---

<sup>1</sup>This point is sometimes obscured within revealed preference approaches where completeness of the revealed preference relation constructed from an agent's choices can be assured by forcing the agent to make choices. But the fact that an agent is forced to choose between two prospects does not make it irrational that they lack a preference between them - its simply means that they must choose in the absence of supporting preference-based reasons.

for representing attitudinal states to make the questions we raise here relevant.

This paper will attempt to address some of the limitations of Bayesian models by dropping the completeness assumption, drawing on a technique advocated by, amongst others, Isaac Levi, I. J. Good and Richard Jeffrey. Instead of representing the judgemental states of agents by precise measures of their partial attitudes, they will be represented by sets of such precise measures - intuitively the set of permissible opinions. The idea is that the common features of these measures represent what is settled for the agent, the differences between them what is not. Thus if the agent regards the probability of rain tomorrow to be no less than 0.5, then her judgemental state will contain only probability functions assigning values greater than or equal to 0.5 to the prospect of rain tomorrow. With such a representation in hand some basic types of transformation of a states of mind will be characterised, compared with those found in AGM-style models of attitude revision, and employed to describe how agents make up their mind when deliberating.

## 2 States of Mind

The discussion of indeterminate or incomplete attitudes will be framed here by the adoption of the framework of Jeffrey-Bolker decision theory in which an agent's judgemental state, or state of mind, is given by her degrees of belief and desire for a set of prospects (or, equivalently, the propositions identifying them). The underlying set of prospects  $\Omega = \{A, B, \dots\}$  is assumed to form a Boolean algebra containing the unit (tautology),  $T$ , but with the zero (contradiction),  $F$ , removed. For any two prospects  $X$  and  $Y$ ,  $X \vee Y$  denotes their join (disjunction),  $XY$  their meet (conjunction) and  $\neg X$  the complement (negation) of  $X$ . We initially assume that the set of prospects is static and that what changes is the agent's attitudes to them, but this assumption is dropped later on.

The agent's preferences are modelled as a two place relation on  $\Omega$  which is assumed to be transitive, but not necessarily complete. Throughout  $\succeq$  will denote the agent's prior preferences and  $\succeq^*$  her posterior ones; posterior that is to whatever experience prompts the change in her attitudes. The relations of strict preference,  $\succ$ , and indifference,  $\approx$ , between prospects are related to  $\succeq$  in the usual way. The agent's preferences, and their evolution over time, are the dependent variables in this discussion - the features of the agent that require explanation or rationalisation by her state of mind and changes in it.

The states of mind of the ideally rational and maximally opinionated agents of Bayesian theory are standardly represented by a *single* pair of functions: a probability and desirability (or utility<sup>2</sup>) measuring, respectively, their degrees of partial belief and desire for relevant prospects. To allow for less opinionated

---

<sup>2</sup>Desirability is expectation of utility and is thus determined by a specification of probability and utility. I choose to work with desirabilities defined over the entire set of prospects rather than with a utility measure restricted to a special class of them (the 'consequences') - which is the more common way of doing things - because it allows for a unified treatment of degrees of belief and desire.

agents, states of mind are instead represented here by a *set* of such functions, each of which may be thought of as a permissible sharpening of the agent's actual beliefs and desires. Such sharpenings, or rather equivalence classes of them, will be termed avatars of the agent, so that the pairs of probability and desirability functions making up an agent's state of mind can be interpreted as measures of the agent's avatar's degrees of belief and desires.<sup>3</sup>

Formally, a state of mind of a rational agent is a set  $S = \{ \langle p_j, v_j \rangle \}$  of pairs of real-valued functions, respectively on  $\Omega$  and  $\Omega - \{F\}$ , where each  $p_j$  is a probability measure and each  $v_j$  a desirability measure satisfying:

1. *Normality*:  $v_j(T) = 0$
2. *Averaging*: If  $XY = F$  and  $p_j(X \vee Y) \neq 0$ , then:

$$v_j(X \vee Y) = \frac{v_j(X) \cdot p_j(X) + v_j(Y) \cdot p_j(Y)}{p_j(X) + p_j(Y)}$$

Each of the pairs of functions  $\langle p_j, v_j \rangle \in S$  also provide a basis for representing an agent's avatar's conditional degrees of belief and desire - her degrees of belief and desire under the hypothesis that some or other condition holds. This is achieved via the definitions:

**Conditional Probability** If  $p_j(A) > 0$ , then:

$$p_j(X|A) := \frac{p_j(XA)}{p_j(A)}$$

**Conditional Desirability** If  $p_j(A) > 0$ , then:

$$v_j(X|A) := v_j(XA) - v_j(A)$$

No further assumptions will be made in this paper about the structure of states of mind. But in some circumstances, or rather under certain interpretations of the set-of-functions model of states of mind, it is reasonable to require of states of mind that they satisfy closure conditions of one kind or another. For instance, if the agent is not fully opinionated about whether it will rain tomorrow because she holds limited information about its prospects, then it seems natural that if she regards both 0.5 and 0.7, say, as permissible probabilities for rain, then all real numbers between 0.5 and 0.7 should be permissible probabilities.<sup>4</sup> Similarly, if the agent cannot determine whether a good is worth £5 or £7 pounds, then arguably she should regard £6 as a permissible price too. A number of different formal conditions could be imposed to express these intuitive requirements; for instance:

---

<sup>3</sup>I use the word avatar here because each sharpening is an opinionated instantiation of the agent's attitudes - her opinionated ways of being, so to speak.

<sup>4</sup>Levi [16], for instance, imposes this requirement.

**Condition 1** *Belief Convexity:* If  $\langle p_1, v \rangle, \langle p_2, v \rangle \in S$  then for  $0 \leq \alpha \leq 1$ :

$$\langle \alpha p_1 + (1 - \alpha) p_2, v \rangle \in S$$

**Condition 2** *Utility Convexity:* If  $\langle p, v_1 \rangle, \langle p, v_2 \rangle \in S$  then for  $0 \leq \beta \leq 1$ :

$$\langle p, \beta v_1 + (1 - \beta) v_2 \rangle \in S$$

**Condition 3** *Joint Convexity:* If  $\langle p_1, v_1 \rangle, \langle p_2, v_2 \rangle \in S$  then for  $0 \leq \alpha, \beta \leq 1$ :

$$\langle \alpha p_1 + (1 - \alpha) p_2, \beta v_1 + (1 - \beta) v_2 \rangle \in S$$

Belief convexity is required if the agent's partial beliefs are to be represented by probability intervals, utility convexity if her partial desires are to be represented by utility ones. However, the last of the three conditions - joint convexity - cannot be satisfied by an agent in a consistent state of mind. For instance, consider the pairs  $\langle p_1, v_1 \rangle$  and  $\langle p_2, v_2 \rangle$  such that  $p_1(X) = 0.6$ ,  $p_2(X) = 0.4$ ,  $v_1(X) = -4$ ,  $v_1(\neg X) = 6$  and  $v_2(X) = 0 = v_2(\neg X)$ . Let  $p = \alpha p_1 + (1 - \alpha) p_2$  and  $v = \beta v_1 + (1 - \beta) v_2$ . Then  $p(X) = 0.4 + 0.2\alpha$ ,  $p(\neg X) = 0.6 - 0.2\alpha$ ,  $v(X) = \beta v_1(X) + (1 - \beta) v_2(X) = -4\beta$  and  $v(\neg X) = \beta v_1(\neg X) + (1 - \beta) v_2(\neg X) = 6\beta$ . By the axioms of averaging and normality:

$$\begin{aligned} v(T) &= -4\beta p(X) + 6\beta p(\neg X) \\ &= -4\beta(0.4 + 0.2\alpha) + 6\beta(0.6 - 0.2\alpha) \\ &= 2\beta - 2\alpha\beta = 0 \end{aligned}$$

Hence  $\beta = 0$  or  $\alpha = 1$  and  $\langle p, v \rangle$  must be either  $\langle p_1, v_1 \rangle$  or  $\langle p_2, v_2 \rangle$ , in contradiction with the joint convexity condition.

## 2.1 Interpretations

Neither the idea that different partial attitudes are permissible, nor the use of sets of measures to represent indeterminacy or incompleteness in agents' states of mind is new (though discussion has largely been centred on probabilistic indeterminacy). The closely related ideas of probability intervals, upper and lower probabilities and comparative probabilities have a long history going back to at least Keynes, Koopman and Borel (see [24]), while use of sets of measures is advocated by, amongst others, Levi [14],[15],[16], Jeffrey [12], Good [8], Voorbraak [22] and Nehring [19]. But the literature contains quite a wide range of applications of the idea and it is useful to get some clarity on this. Broadly speaking there are three different kinds of interpretation that can be given to the sets of measures formalism - namely, subjective incompleteness, measurement imprecision and objective indeterminacy - corresponding to cases in which an agent fails to reach a judgement, cases in which the evidence underdetermines our representation of their judgemental state, and cases in which there is no fact of the matter as to what judgement the agent has reached.

- A *Subjective incompleteness*: The agent hasn't, won't or can't make up her mind because she lacks the means or motivation to do so (at least temporarily). This is doubtless the state most of us find ourselves in with regard to a good number of prospects. Often with good reason: when we are faced with inadequate or conflicting evidence or advice, or the costs of making an informed judgement exceed the benefits, it may be better if we don't reach a judgement. So the incompleteness may be rationally permissible.
- B *Limits of Measurement*: Even in circumstances when rational agents either have or should have complete partial attitudes, these cannot always be precisely measured because of practical limitations on psychological measurement. Use of probability intervals have sometimes been motivated by the possibility of such measurement imprecision, for instance (see Good [8]).
- C *Objective Indeterminacy*: On some views, it is not merely the case that the evidence we hold about agents underdetermines our measurement of their mental states, it is that there is no fact of the matter concerning whether they are in one or another attitudinal state. Consequently, their states of mind cannot, even in principle, be represented by a single probability-desirability pair. Richard Jeffrey [12], for instance, regarded the fact that in his framework preferences do not uniquely determine representations of the agent's degrees of belief and desire to be a reflection of the fact that the two are essentially entangled in practical judgements about what to do, rather than as a sign that for measurement purposes preferences need to be supplemented with additional evidence.

These interpretations are not mutually exclusive of course and it would be possible to use the sets of measures formalism in circumstances in which imprecision derives in part from the agent, in part from the evidence and in part from the facts. As our earlier discussion suggests, however, it is subjective indeterminacy that is the main focus of our interest and hereafter we will ignore the other two. But even within this broad domain there are a number of different interpretations that can be given to the idea that different partial attitudes are permissible.

- A1** *Unawareness*: The agent may not have taken an attitude to some prospect simply because she is unaware of its existence. So, for instance, we have no attitude to people we have not heard about, to restaurants in the neighbourhood that have yet to open, or to products that have yet to be marketed. The possible resulting disparity between what the modeller and the agent is aware of can be resolved by taking the underlying prospect space to be set of things of which the modeller is aware, and by treating every attitude that the agent might take to the objects of which she is unaware (consistent with her attitudes to other prospects) as permissible.

Each avatar of the agent then represents a permissible extension of her attitudes to the objects of which she is unaware.<sup>5</sup>

- A2** *Boundedness.* The agent may simply not have made up her mind about a particular question because she has either not thought it through completely or has deliberately suspended judgement. This could be either because she is awaiting further information or because the issue is of insufficient importance to justify expenditure of the deliberational resources required to reach a judgement. For instance, she may have decided that she will take a holiday by the sea, but not where yet, intending to do some reading about her options first. Then each of her avatars represents a possible opinion on where would be best to go.
- A3** *Ignorance:* The agent may be unable to arrive at a judgement because she lacks the informational basis for doing so. This seems to be the kind of situation in which subjects find themselves, for instance, when placed in an Ellsberg paradox set-up in which the consequences of their decisions depend on the colour of a ball drawn from an urn containing an unknown proportion of balls of different colours. Many authors argue that in these kinds of situation the agent is not merely in a state of uncertainty in the sense that they don't know for sure which colour ball will be drawn but can assign a probability to the prospect of each colour, but are rather in a state of ignorance in the sense that, such are the limits on what they know and can find out, that they have no non-arbitrary basis for assigning such a probability. See for instance, Voorbraak [22], Levi [16], Steele [21] and Nehring [18], [19].
- A4** *Conflict:* The agent may be in a state of unresolved conflict which prevents her from reaching a single judgement on an issue. This can derive, for instance, from conflicting epistemic or value commitments, an issue on which Isaac Levi has written eloquently (see [16]), or because the agent is not psychologically unified, as may be the case in group decision making when the group lacks cohesion or in intertemporal decision making when the agent's tastes change over time. Each avatar here represents an opposed viewpoint, one which remains open not simply because the agent has not got around to resolving the question of which viewpoint is to be hers, but because she cannot find a way of resolving it and must make her choices in the light of the conflict.

These cases may well lie on a spectrum running from those in which an agent has not finished making up her mind, but has at her disposal the means to do so (e.g. she has all the necessary information, is able to make the correct inferences, has the time and energy to draw them, etc.), through those in which there are temporary obstacles to her reaching a judgement (lack of awareness, information or inferential tools), all the way to those cases in which she cannot

---

<sup>5</sup>This is obviously not the only way in which unawareness can be modelled. Indeed in the final section I will suggest a different way.

make up her mind because of incompatible commitments or insurmountable obstacles to acquiring or processing relevant information (e.g. because of the complexity of the issue or intractability of the conflict).

## 2.2 Rationalising Preference

Any pair of probability and desirability functions induces a ranking of the set of prospects by desirability value. If the agent is maximally opinionated and her preferences for prospects go by their desirability, then the induced desirability ranking may be thought of as explaining or rationalising the agent's preferences. Less opinionated agents' preferences should also cohere with their desirability judgements. In this more general case, a state of mind *explains* or *rationalises* an agent's preferences whenever every pair of probability and desirability functions making it up implies a ranking of the prospects that is consistent with the agent's preferences. More formally, we say:

**Definition 4** A state of mind  $S = \{ \langle p_j, v_j \rangle \}$  *explains* or *rationalises* the preference relation  $\succeq$  iff, for all  $X, Y$  in the domain of  $\succeq$ , it is the case that, for all  $\langle p_j, v_j \rangle \in S$ :

$$X \succeq Y \Rightarrow v_j(X) \geq v_j(Y)$$

Clearly an agent's preferences could be rationalised by more than one state of mind, since any subset of a state of mind will rationalise the preferences rationalised by the state of mind itself. (However only one rationalising state of mind contains *all* pairs of probability and desirability measures jointly consistent with her preferences. Such a state we call *maximal*.) The more complete an agent's preferences, the smaller the set of pairs of functions rationalising them. But even complete preferences can be rationalised by more than one state of mind, because the same desirability ordering can be induced by different desirability functions. Where the differences in values assigned to prospects by two desirability functions are arbitrary - because, for instance, they are attributable to nothing more than different choices of scale - these functions may be regarded as substantially equivalent. Similarly, two states of mind may be regarded as substantially equivalent if each of the functions belonging to one is equivalent to a function belonging to the other, and vice versa.

The claim that different states of mind may be substantially equivalent is consistent with different views as to what makes for such equivalence. There is a long tradition in decision theory of regarding any differences in measures of the agent's attitudes not derivable from features of the agent's preferences as arbitrary scaling features and without empirical content, a view that I will adopt here (though not very much depends on it). From this point of view, the significance of the foundational representation theorems of decision theory lies in the fact that they establish conditions for the existence of explanatory states of mind as well as the relationships between the elements of them by establishing which transforms of a given explanatory state of mind also explain the agent's preferences. The precise transforms allowed under different axiomatisations of preference need not concern us here however.

On this position, one has an opinion on a prospect iff one is able to place it in a preference ordering of prospects. Substantial differences of opinion should be reflected in differences in this ordering. In this spirit:

**Definition 5** *A state of mind  $S$  is said to be **opinionated** whenever, for all  $X, Y \in \Omega$ , it is the case that  $\forall \langle p_j, v_j \rangle \in S, v_j(X) \geq v_j(Y)$  or  $\forall \langle p_j, v_j \rangle \in S, v_j(Y) \geq v_j(X)$  and non-opinionated otherwise. An avatar of an agent is an opinionated subset of her state of mind.*

It follows immediately from these definitions that:

1. If  $S$  is an avatar and  $\langle p_j, v_j \rangle, \langle p_k, v_k \rangle \in S$  then

$$v_j(X) \geq v_j(Y) \Leftrightarrow v_k(X) \geq v_k(Y)$$

2. If  $S$  is an avatar that explains  $\succeq$  then  $\forall \langle p_j, v_j \rangle \in S, X \succeq Y \Leftrightarrow v_j(X) \geq v_j(Y)$
3. If  $\succeq$  is complete then  $S$  explains  $\succeq$  iff  $S$  is opinionated.

### 3 Revising Opinionated States of Mind

Agents' judgemental states can change in a number of different ways. They can revise their opinions about one or more prospects, they can form new opinions and they can withdraw or suspend opinions currently held. In AGM theory these changes are termed revision, expansion and contraction; here I will speak of attitude/judgement/opinion *revision*, *formation* and *withdrawal* since these terms are less closely tied to the choice of formal framework for modelling the changes. Hansson [9] notes that agents may also extend or restrict the domain over which they take attitudes, operations which he calls addition and subtraction, but which will be termed opinion *extension* and *restriction* here. Our aim in the next few sections is to define formal operations on states of mind, construed as sets of pairs of probability and desirability measures, corresponding to each of these ways in which agents can change their opinions.

My starting point is the Bayesian conditioning model or rather the family of Bayesian conditioning models designed to handle different types of experience. Like the AGM theory, Bayesianism provides 'perturbation-propagation' models in which the change in an agent's state of mind is viewed as a two-step process. In the first stage, the agent changes an attitude to a particular prospect or, more generally, some set of prospects. In the second, she adjusts her attitudes to all other possibilities in order to restore consistency. The processes inducing the initial change are not themselves modelled: they might include sensory experience, deliberation, reception of a message from a reliable information source, or even hypnosis.

The most well-known theory belonging to the Bayesian family is that of classical conditioning according which agents should update their belief upon

receipt of the information that X, by adopting as their new degrees of belief their old conditional degrees of belief, given X. This way of updating is demonstrably the uniquely rational way of revising belief in response to this information just in case the perturbatory experience leaves the agent with posterior conditional degrees of belief, given X, identical to her prior ones i.e. just in case the perturbatory effect of experience is completely described by the constraint that her new degrees of belief in X equal one. This condition is not always satisfied: there can be no such prospect X such that the effect of experience is exhaustively described in terms of the agent coming to believe X with degree one. This limitation is largely overcome in Jeffrey conditioning and generalised conditioning models (see Jeffrey [13], Bradley [4]), which prescribe a particular kind of belief and desire change in cases where the perturbatory effects of experience can be localised some particular partition of the space of prospects. In fact *all* changes in belief and desire can be modelled in these terms, at least when the algebra of prospects is countable.<sup>6</sup>

The essence of this theory of attitude change can be stated as follows. Suppose that the effect of experience is exhaustively described by a redistribution  $\alpha\beta = \{\alpha_i, \beta_i\}$  of probability and desirability across a particular partition  $A = \{A_i\}$  of the space of prospects, with each  $A_i$  being assigned probability  $\alpha_i$  and desirability  $\beta_i$ . Then an agent revises her state of mind by **generalised conditioning** on the new values for the  $A_i$  just in case her new state of mind  $\langle p_{\alpha\beta}^*, v_{\alpha\beta}^* \rangle$  is related to prior state  $\langle p, v \rangle$  by, for all prospects  $X \in \Omega$  :

$$p_{\alpha\beta}^*(X) = \sum_i p(X|A_i) \cdot \alpha_i \quad (1)$$

$$v_{\alpha\beta}^*(X) = \sum_i [v(X|A_i) + \beta_i] \cdot p_{\alpha\beta}^*(A_i|X) \quad (2)$$

Note that  $p_{\alpha\beta}^*(X)$  and  $v_{\alpha\beta}^*(X)$  are defined only when the  $p(X|A_i)$  and  $v(X|A_i)$  are. On the standard definitions given before, conditional probabilities and desirabilities are not defined for conditions of probability zero, which means that  $p_{\alpha\beta}^*(X)$  and  $v_{\alpha\beta}^*(X)$  are not defined for any partition  $A$  containing an element of zero probability. This may be of little practical concern, but it is sometimes more technically convenient to have generalised conditioning defined for all partitions. To obtain this it would suffice to work with Popper-Renyi conditional probability functions which are defined even for cases in which the unconditional probability of a condition (now defined as its conditional probability given the truth of the necessary prospect) is zero.<sup>7</sup> This in turn would require us to characterise agents state of mind in terms of conditional probabilities and desirabilities, rather than unconditional ones, but we shall not pursue that path further here.

Equation 1 is the familiar expression for Jeffrey conditioning, equation 2 the extension of this theory to desirabilities. In the limiting case when the

<sup>6</sup> See Bradley [5], Theorem 5, for an exact formulation and proof of this claim.

<sup>7</sup> For further detail, see Spohn [20].

probabilities of one of the elements of the relevant partition go all the way to one, generalised conditioning just reduces to classical Bayesian conditioning in which the agent adopts her prior conditional attitudes as her posterior attitudes.

Generalised conditioning on the partition  $A$  is demonstrably the uniquely rational way of changing one's attitudes in response to experience whenever one's posterior conditional desirabilities, given the elements of the partition in question, are identical to one's prior conditional desirabilities given these prospects, i.e. whenever the following Rigidity condition holds:

**Condition 6** *Rigidity with respect to the  $A_i$ :  $v_{\alpha\beta}^*(\cdot|A_i) = v(\cdot|A_i)$*

This requirement - that conditional attitudes, given the elements of relevant partition, be invariant under changes to unconditional attitudes to these elements - is what I take to be defining characteristic of the Bayesian theory of attitude revision.<sup>8</sup> The theory is, in sense, a tautology for if the perturbatory effect of experience is correctly and exhaustively described by a redistribution of probability and desirability over a particular partition, then Rigidity with respect to that partition must hold. For if it failed for some partition element  $A_i \in A$  and prospect  $X$ , for instance, then it would follow that the desirability of  $AX$  relative to that of  $A\neg X$  had changed, which would imply in turn that the perturbatory effect of experience was not localised to the partition  $A$ . But in a deeper sense it expresses the view that the effect of experience can (usually) be localised to a particular partition in this way, because the evidence that experience provides for or against the truth of any element of a partition will not in itself provide the agent with reason to change her attitudes to other prospects given its truth.

It does not follow of course that generalised conditioning offers a complete theory of attitude change.<sup>9</sup> Two questions in particular still need to be addressed: firstly how generalised conditioning works when agents have non-opinionated prior states of mind, and, secondly, whether conditioning is the only way of changing one's mind. This second question can be profitably addressed by considering how conditioning relates to the various types of attitude change countenanced by AGM theories and conversely what (other) ways of transforming quantitative degrees of belief and desire might correspond to operations like expansion, revision and contraction.

Early consideration of these issues was given by Gärdenfors [6], who offered a set of postulates for probabilistic revision, expansion and contraction. He also argued that classical Bayesian conditioning should be regarded as the probabilistic equivalent of belief expansion because it preserves beliefs of probability one (in contrast, for instance, to probability change by imaging). His characterisations of these probabilistic operations were confined however to the special case in which the perturbatory experience induces a new full belief in some prospect.

<sup>8</sup>This is perhaps a little imperialistic as there are approaches that call themselves Bayesian (e.g. that of Levi [15]) which lay little store by the rigidity conditions. But classification is an inexact science.

<sup>9</sup>Nor do I claim that it is the only interesting way of generalising classical Bayesianism. Levi [14],[15], for instance, does so in a different way.

So, for instance, he gave no consideration to the question of the status of Jeffrey conditioning.

More recently, Frans Voorbraak [23] has argued that Bayesian conditioning is not a form of probabilistic expansion at all (although classical conditioning bares certain similarities to it) and, indeed, that it involves a type of attitude change different from any of the AGM operations. Probabilistic expansion on his account is achieved by an operation called constraining and which is defined on sets of probability functions rather than single ones. I fully concur with this latter claim of his, and will offer an extension of it below, but disagree with both him and Gärdenfors on the status of Bayesian conditioning. I think that conditioning is best regarded as a genuine *revision* operation which, in its limiting cases (those described by classical conditioning), has features similar to expansion. After all conditioning does require one to revise one's attitudes i.e. to give up one's current ones and adopt new ones.

Resistance to this classification derives, it would seem, from that the fact that there is a sense in which conditioning does not require agents to give up anything that they already believe and that instead of revising their opinions Bayesian agents 'merely' update their attitudes in the light of new information. This impression is mistaken, however. Jeffrey conditioning, and its desirabilistic correlate, is not restricted to cases of updating with new information, but is also applicable to situations in which an agent reconsiders her opinions in the light of fresh evidence. For instance, having spotted what I take to be my friend walking on the other side of the road I form a very high degree of belief in the prospect of him being in the neighbourhood. If however I receive a phone call from him shortly thereafter in which he claims to be in the Bahamas, I am likely to revise this opinion considerably. In other words, Jeffrey conditioning is reversible in a way that classical conditioning is not - in my opinion, an essential characteristic of a revision operation.

## 4 Changes to Incomplete Attitudes

### 4.1 Revision

Formally, the perturbation-propagation models of attitude change we have been discussing are mappings from a prior attitude state to a posterior one, as a function of the constraints imposed by experience, or rather of the constraints the agent adopts on her attitudes as a result of experience. The generalised conditioning model for opinionated states of mind, for instance, maps a prior state of mind  $S$  to a posterior state of mind  $S_{\alpha\beta}^* = \langle p_{\alpha\beta}^*, v_{\alpha\beta}^* \rangle$ , as a function of the redistribution  $\alpha\beta$  of probability and desirability across some partition  $A$ . In the light of the previous discussion, I will call  $S_{\alpha\beta}^*$  the state of mind obtained from a prior state  $S$  by *revision* on the partition  $A$ .

Extending this conditioning model of revision to the states of mind of non-opinionated agents is fairly straightforward and is most naturally achieved by taking a posterior state of mind to be simply the union of the probability-

desirability pairs obtained by generalised conditioning on the elements of the prior state of mind. More formally, we extend as follows the definition of generalised conditioning to revisions of non-opinionated states of mind prompted by a redistribution of probability and desirability across a partition  $\{A_i\}$ :

$$S_{\alpha\beta}^* = \{\langle p_{\alpha\beta}^*, v_{\alpha\beta}^* \rangle : \langle p, v \rangle \in S \text{ and } \forall A_i \in A, p(\cdot|A_i) \text{ and } v(\cdot|A_i) \text{ are defined}\}$$

This definition leaves open the question of whether conditional probabilities and desirabilities are defined for zero-probability prospects. If they are, then we have a pure revision model. If not, then in cases where experience induces an allocation of probability or desirability to zero-probability events, revision will coincide with the opinion formation operation defined below.

Before turning to other kinds of changes to non-opinionated states of mind, it is worth noting that revision of non-opinionated states of mind can lead to an agent acquiring new preferences and to her withdrawing some. It is even possible that revision of a non-opinionated state of mind can lead to the agent developing a complete set of preferences. Consider, for example, the actions  $A_1$  and  $A_2$  with consequences, conditional on the event  $E$ , having the desirabilities given in the following table:

	$E$	$\neg E$
$A_1$	1	0
$A_2$	0	2

Suppose that the agent has two avatars  $\langle p_1, v_1 \rangle$  and  $\langle p_2, v_2 \rangle$ , which are in disagreement on the probability of  $E$ . Then by definition 4, for any avatar  $i \in \{1, 2\}$ :

$$A_1 \succeq A_2 \Leftrightarrow p_i(E|A_1) \geq 2p_i(\neg E|A_2)$$

For simplicity consider the case where  $E$  is probabilistically independent of both actions. Then we have the following cases:

1.  $p_i(E) \geq 2/3$        $A_1 \succeq A_2$
2.  $p_i(E) \leq 2/3$        $A_2 \succeq A_1$
3.      Else       $A_2 \not\succeq A_1, A_2 \not\succeq A_1$

A preference gain will occur whenever the initial situation is case 3 and conditioning on new information drives the probabilities of both avatars into cases 1 or 2. On the other hand a preference loss will occur whenever conditioning drives the probabilities of one of the avatars from case 1 or 2 into case 3. For example, suppose that the agent observes that  $O$ . Then:

*Preference gain:* Suppose that initially  $p_1(E) = 1$ ,  $p_2(E) = 1/2 = p_2(O)$  and  $p_2(EO) = 1/3$ . Then if both avatars condition on observation  $O$ , it will follow that  $p_1^*(E) = 1$  and  $p_2^*(E) = 2/3$ . Hence  $A_1 \succeq^* A_2$ .

*Preference loss:* Suppose that initially  $p_1(E) = 1$ ,  $p_2(E) = 3/4 = p_2(O)$  and  $p_2(EO) = 1/4$ . Then if both condition on observation  $O$ ,  $p_1^*(E) = 1$  and  $p_2^*(E) = 1/3$ . Hence  $A_1 \not\succeq^* A_2$ .

## 4.2 Opinion Formation

Generalised conditioning is a technique for revising mental attitudes, both complete and incomplete. But this is not the only way in which our state of mind

changes. Sometimes we make up our mind about something and come to hold an opinion where once we had none. Equally we sometimes withdraw or suspend judgements, thereby giving up opinions, such as when our reasons for them are undermined or when we want to create space for deliberating with someone who holds different opinions or when we want to reason hypothetically about possibilities ruled out by our current judgements. These actions are different from revision: in essence revision is a passage from one set of permissible judgements to another set, while opinion formation and withdrawal involve respectively a reduction in the set of permissible judgements and an expansion of them.

Consider firstly the case in which an agent makes up her mind about some question. She might, for instance, learn that some proposition is true or come to fully adopt a definite probability for some event on the basis of statistical evidence, or perhaps simply adopt some weaker constraint on her opinions. Whatever the judgement, it can be represented by the set of pairs of probability and desirability functions consistent with it. For instance if the agent comes to believe that  $X$ , then this judgement is represented by the set of all probability functions assigning probability one to  $X$  (paired with any desirability function consistent with it). If she comes to believe that the probability of  $X$  lies in the interval between 0.5 and 0.7, then her judgement is represented by all probability functions assigning a probability in this interval to  $X$ , and so on. Call this the constraint set and denote it by  $C$ . Then we say that the agent changes her state of mind  $S$  by **opinion formation** from  $C$  just in case her new state of mind,  $S_C^+$ , is the set  $S \cap C$ .

In line with the treatment of revision, and without significant loss of generality, we focus attention here on the case where the new constraint takes the form of an assignment  $\alpha\beta = \{\alpha_i, \beta_i\}$  of probabilities and desirabilities across a partition  $A = \{A_i\}$ , with each  $A_i$  being assigned probability  $\alpha_i$  and desirability  $\beta_i$ . In this case:

$$S_{\alpha\beta}^+ = \{\langle p, v \rangle \in S : p(A_i) = \alpha_i, v(A_i) = \beta_i\}$$

In the special case where the constraint on new attitudes derives from the information that some particular prospect  $A$  is true, then  $S_A^+ = \{\langle p, v \rangle \in S : \langle p(\cdot|A), v(\cdot|A) \rangle \in S\}$ , which is just the set of permissible prior probability functions assigning probability one to  $A$ .<sup>10</sup>

### 4.3 Opinion Withdrawal

The last case to consider is one in which an agent wishes to suspend or withdraw judgement on some issue for one of the reasons mentioned before. In suspending judgement she *reduces* the constraints on her attitudes, so the state of mind that she achieves should be less opinionated than her original one i.e. her new state of mind should be a superset of her initial one. How would a Bayesian

<sup>10</sup>The kind of belief change associated with opinion formation is termed probabilistic constricting by Voorbraak [23], an operation that he considers to be the true probabilistic analogue of belief expansion in the AGM framework.

agent form such a superset? Our guide here are the considerations underlying the rigidity assumption characteristic of Bayesian revision; namely that the grounds we have for judgement on a prospect  $A$  are independent of the grounds we have for our conditional judgements given  $A$ . Translated into the context of suspension of judgement this Bayesian dictum yields the requirement that conditional attitudes, given some prospect, be unaffected by the suspension of attitude towards that prospect. Formally, let  $S = \langle p, v \rangle$  be an opinionated state of mind and  $S_{\{A_i\}}^-$  be the state of mind reached by suspension of attitude towards the  $A_i$ . Then we require:

**Condition 7** *Rigidity with respect to the  $A_i$ : If  $v_j^- \in S_{\{A_i\}}^-$ , then  $v_j^-(\cdot|A_i) = v(\cdot|A_i)$  whenever both the  $v_j^-(\cdot|A_i)$  and  $v(\cdot|A_i)$  are defined.*

Rigidity imposes a constraint on the amount of change in a state of mind that can be induced by the withdrawal of opinion on a partition of prospect space. But it alone does not suffice to uniquely determine the outcome of such a process. Consider a case, for instance, in which suspension of opinion is motivated by the discovery that someone that one respects has a opinion different to one's own on some particular issue. Should one adopt as permissible just the union of the two states of mind, all convex combinations of the two, or a state containing all possible attitudes to the question under dispute? A principle of minimum change would favour the former, but in cases where one thinks that both parties have part of the truth, perhaps because they hold different incomplete evidence, then the second path seems preferable. Even this may be too restrictive however, since the fact that my evidence and your evidence both point to, say, a low probability for some prospect, does not imply that our evidence combined would not point to a high probability for the prospect. Complete suspension of opinion, the third strategy, is favoured by the thought that this possibility should not be precluded until the grounds for the difference of opinion have been clarified.

Here we will pursue the third strategy and describe a method of suspending judgement on a partition in which an agent enlarges her set of permissible avatars so as to include all those avatars with conditional attitudes given the elements of the partition that are identical to a prior avatar of hers, but with any consistent set of attitudes to the partition elements themselves. We start by characterising suspension or withdrawal of judgement in the case of an opinionated agent. Let  $\langle p_j, v \rangle$  be the agent's prior opinionated state of mind and suppose that she suspends judgement of a partition  $\{A_i\}$  such that  $p(A_i) > 0$  for all elements  $A_i$  of the partition. Then the agent changes her mind by (complete) **opinion withdrawal** on the partition  $\{A_i\}$  just in case for all  $\alpha = \{\alpha_i\}$ , where  $0 \leq \alpha_i \leq 1$ , and all  $\beta = \{\beta_i\}$  where  $\beta_i \in \mathfrak{R}$ , the pair of probability and desirability functions  $\langle p_j^{-\alpha}, v_j^{-\beta} \rangle$  belongs to posterior state of mind, where  $\forall X \in \Omega$ :

$$\begin{aligned} p_j^{-\alpha}(X) &= \sum_i \alpha_i \cdot p_j(X|A_i) \\ v_j^{-\beta}(X) &= \sum_i [v(X|A_i) + \beta_i] \cdot p_j^{-\alpha}(A_i|X) \end{aligned}$$

Extending this model of attitude suspension to the states of mind of non-opinionated agents is fairly straightforward and is most naturally achieved by taking the posterior (post-withdrawal) state of mind to be simply the union of the probability-desirability pairs obtained by opinion withdrawal on the elements of the prior state of mind. More formally, we extend as follows the definition of opinion withdrawal on a partition  $\{A_i\}$  to non-opinionated states of mind:

$$S_{\{A_i\}}^- = \{\langle p_j^{-\alpha}, v_j^{-\beta} \rangle : \langle p_j, v_j \rangle \in S \text{ and } \forall A_i \in A, p_j(\cdot|A_i) \text{ and } v_j(\cdot|A_i) \text{ are defined}\}$$

Given our definitions of conditional probability and desirability, this definition of opinion withdrawal in effect forbids an agent from suspending her opinion on a partition when her prior probability for some element of it is zero. As in the case of opinion revision this limitation could be overcome by working with measures of conditional attitudes that are defined for zero-probability prospects.

The state of mind achieved by opinion withdrawal satisfies the Rigidity requirement. It is also the least opinionated state of mind to do so, so that any state of mind achieved by a less complete withdrawal of opinion can be recovered by addition of further constraints on opinion. Both claims are proved in the appendix as Theorem 10.

#### 4.4 Connections

Forming an opinion and withdrawing are operations that are, on the face of it, quite different from revision. Revision is essentially an operation on permissible partial attitudes, while formation and withdrawal affect the number of attitudes that are permissible. This is exemplified by the fact that any probability-desirability pair  $\langle p, v \rangle$  belongs to the posterior state  $S_{\{A_i\}}^-$  or  $S_{\alpha\beta}^+$  only if it belongs to prior state  $S$ , while this will only be true for revision in degenerate cases.

On the other hand, opinion formation and withdrawal are opposite operations in the sense that withdrawing opinion across a partition and then forming a new opinion on the partition elements using the original assignment of probability and desirability to them, returns one to one's original state of mind, i.e. if  $\alpha\beta = \{\alpha_i, \beta_i\}$  is an assignment of probability and desirability across  $A = \{A_i\}$ , and  $S = \{\langle p, v \rangle\}$  is a prior state of mind such that  $p(A_i) = \alpha_i$  and  $v(A_i) = \beta_i$ , then:

**Proposition 8** (*Recovery*)  $S = (S_{\{A_i\}}^-)_{\alpha\beta}^+$

This is proven in the Appendix as Corollary 12. Note, however that forming an opinion using some assignment of probability and desirability to a particular partition, and then withdrawing one's opinion on that partition, need not return one to one's original state of mind. Indeed it will only do so if the original state of mind was completely non-opinionated with regard to the partition in question.

Despite the difference between revision and opinion formation and withdrawal, there is an interesting relation between them; namely that revising on

a set of new values for the elements of a partition is equivalent to withdrawing one's judgements on this partition and then forming a new judgement based on these new values. Formally, let  $S = \{ \langle p_j, v_j \rangle \}$  be any state of mind and  $\alpha\beta = \{ \alpha_i, \beta_i \}$  an assignment of probability and desirability across the partition  $\{A_i\}$ . Then:

**Proposition 9**  $S_{\alpha\beta}^* = (S_{\{A_i\}}^-)_{\alpha\beta}^+$

(Proof in the Appendix as Theorem 15).

A rather similar equivalence is central to the AGM theory, where it carries the title of the Levi Identity. By the Levi Identity revising a belief set by some sentence  $A$  is equivalent to withdrawing  $\neg A$  from the belief set and then expanding with  $A$ . This suggests that the withdrawal operation characterised here is a good candidate to serve as the counterpart to contraction in the AGM framework. But there are some differences. Another equivalence central to AGM - the Harper Identity - says that the belief set formed by withdrawing a sentence  $A$  is the same as taking the union of the prior belief set and the set formed by revising by  $A$ . But no such equivalence result can be stated in the Bayesian framework sketched here, at least without defining the complement of assignment of probability and desirability across a partition.<sup>11</sup> Another difference: opinion withdrawal is weakly monotonic in the sense that if  $S$  and  $T$  are any two states of mind such that  $S \subseteq T$  then  $S_{\{A_i\}}^- \subseteq T_{\{A_i\}}^-$  (this is an immediate consequence of Lemma 13 in the Appendix), but AGM contraction is not.

## 5 Deliberation and Domain Change

In this last section, we consider a difficulty for the treatment of deliberation within a Bayesian framework that arises from the tension between the requirement of logical omniscience and the possibility that agents lack settled attitudes to all prospects.

Consider, for example, actions  $A_1$  and  $A_2$  with consequences conditional on the event  $E$  as given in the following matrix:

	$E$	$\neg E$
$A_1$	$C_1$	$C_3$
$A_2$	$C_2$	$C_4$

An agent who has probabilities for  $E$  and  $\neg E$  and desirabilities for the  $C_i$ , but not for the actions  $A_1$  and  $A_2$  is in a state of mental disequilibrium in the sense that she has not accepted all the consequences of her current attitudes. How should we represent her state of mind and the deliberative process by which she derives an opinion on the options before her from her opinions regarding

<sup>11</sup>The natural definition of the complement of a particular assignment is the set of all other assignments consistent with the probability and desirability laws. This definition would allow the statement of a Bayesian equivalent to the Harper identity.

their consequences and the probability that they will arise? The problem we face is that the method for representing incomplete states of mind developed in this paper would recommend that her state of mind be characterised by a set of avatars sharing an opinion on the probability of  $E$  and the desirabilities of the  $C_i$ , but differing on the desirability of the actions. But since the former should determine the latter, this requires that all but one of the avatars be in an inconsistent state of mind.

One way of dealing with the problem is simply to drop the requirement of logical omniscience and permit inconsistent avatars. Inconsistent avatars can then be eliminated by application of deliberation rules derived from the probability and desirability calculus. This approach has a number of attractions. For one, there are independent grounds - most saliently descriptive realism - for dropping the usual requirement of logical omniscience. Secondly, this way of proceeding allows for a unified treatment of two kinds of mental failure. The first, which we can call agent inconsistency, occurs when an agent holds determinate, inconsistent attitudes because her judgements regarding the possible actions do not depend in the right kind of way on her probability and desirability judgements regarding the relevant events and consequences. The second, which we can call avatar inconsistency, occurs when the agent lacks an opinion as to the desirability of the actions, despite the fact that her judgements regarding these events and consequences imply one.

In both cases, inconsistency can be removed by the agent as follows:

1. Withdraw her opinions on the actions;
2. Derive the desirability of the actions from the desirabilities of the  $C_i$  and the probability of  $E$ ;
3. Form an opinion on the actions by taking the calculated desirabilities as the constraint set.

Formally if her initial state of mind is  $S$ , then this procedure yields a new state of mind  $S^* = S_C^+(S_{\{A_1, A_2\}}^-)$  where  $C$  is the set of  $\langle p, v \rangle \in S_{\{A_1, A_2\}}^-$  such that

$$\begin{aligned} v(A_1) &= v(C_1).p(E) + v(C_3).p(\neg E) \\ v(A_2) &= v(C_2).p(E) + v(C_4).p(\neg E) \end{aligned}$$

A second possibility is to treat deliberation, not as a means of *revising* attitudes, but as a way *extending* them to new prospects. This way of proceeding allows attitude change through deliberation to be treated on a par with the attitude change brought about by becoming aware of prospects for the first time. What makes it different from revision of previously inconsistent attitudes is that the agent's initial attitudes are defined on a set of prospects not containing those that she is unaware of or has not yet given deliberative consideration to - hence consistency is not threatened.

To model attitude change through changes in the domain of prospects towards which the agent has attitudes, we need to characterise the class of admissible domain changes. Given any domain of prospects  $\Omega$ , let  $\Omega_E^\oplus$ , called the *extension* of  $\Omega$  by  $E$ , be the closure of  $\Omega \cup E$  under the boolean operations.

And let  $\Omega_R^\ominus$ , called the *restriction* of  $\Omega$  by  $R$ , be the largest subset of  $\Omega$  that does not contain  $R$  and which is closed under the boolean operations.<sup>12</sup> Then  $\Omega_R^\ominus \subseteq \Omega - R \subseteq \Omega \subseteq \Omega \cup E \subseteq \Omega_E^\oplus$ .

Let  $\langle p, v \rangle$  be a pair of probability and desirability functions on  $\Omega$ . Then  $\langle p_E^\oplus, v_E^\oplus \rangle$  and  $\langle p_R^\ominus, v_R^\ominus \rangle$  are respectively said to be an extension of  $\langle p, v \rangle$  to  $\Omega_E^\oplus$  and a restriction of  $\langle p, v \rangle$  to  $\Omega_R^\ominus$  just in case  $\langle p_E^\oplus, v_E^\oplus \rangle$  and  $\langle p_R^\ominus, v_R^\ominus \rangle$  are pairs of probability and desirability functions, respectively on  $\Omega_E^\oplus$  and on  $\Omega_R^\ominus$ , that agree with  $p$  and  $v$  on  $\Omega$ .

Let  $S = \{\langle p_j, v_j \rangle\}$  be a state of mind explaining preferences  $\succeq$  on  $\Omega$ . Then  $S_E^\oplus$ , the extension of  $S$  to  $\Omega_E^\oplus$ , is defined as the set of extensions of the agent's probabilities and desirabilities to the elements of  $S$ , i.e.

$$S_E^\oplus := \{\langle (p_j)_{E_j}^\oplus, (v_j)_{E_j}^\oplus \rangle : \langle (p_j)_{E_j}^\oplus, (v_j)_{E_j}^\oplus \rangle \text{ is an extension of } \langle p_j, v_j \rangle\}$$

Similarly  $S_R^\ominus$ , the restriction of  $S$  to  $\Omega_R^\ominus$ , is defined as the set of restrictions of the agent's probabilities and desirabilities to the elements of  $S$ , i.e

$$S_R^\ominus := \{\langle (p_j)_{R_j}^\ominus, (v_j)_{R_j}^\ominus \rangle : \langle (p_j)_{R_j}^\ominus, (v_j)_{R_j}^\ominus \rangle \text{ is a restriction of } \langle p_j, v_j \rangle\}$$

When an agent derives the deductive consequences of her attitudes, the state of mind she adopts is the extension of her prior attitudes to prospects in some restricted domain to the 'newly considered' prospects. In our example deliberation by domain manipulation can be spelled out as follows:

1. Opinion is first restricted to the Boolean algebra based on  $\{E, \{C_i\}\}$ .
2. Opinion is then extended to the Boolean algebra of prospects based on the extension of the initial algebra to one containing  $\{A_1, A_2\}$ .

Formally if the agent's initial state of mind is  $S = \{\langle p_j, v_j \rangle\}$ , then her post-deliberation state of mind is  $(S_{\{E, \{C_i\}\}}^\ominus)_{\{A_1, A_2\}}^\oplus$ . Suppose that the agent has determinate attitudes to  $E$  and the  $C_i$ . Then, ignoring non-relevant prospects, the state of mind  $S_{\{E, \{C_i\}\}}^\ominus$  must be an opinionated one. This in turn implies that  $(S_{\{E, \{C_i\}\}}^\ominus)_{\{A_1, A_2\}}^\oplus$  must be opinionated since there is a unique extension of her attitudes to the actions consistent with the desirability axioms. Thus her post-deliberative state should be both opinionated and consistent.

## 6 Concluding Remarks

In this paper I have tried to develop a Bayesian theory of attitude change appropriate to non-opinionated, but rational, agents; one that describes not only how incomplete attitudes should be revised, but also how they should be formed and withdrawn i.e. how an agent might make her attitudes more or less determinate. This project takes Bayesian theory into terrain already occupied by AGM style theories of attitude revision and a second objective of this paper has been to compare and contrast the two approaches.

<sup>12</sup>There is a difficulty here: the condition may not pick out a unique set. In this case we let  $\Omega_R^\ominus$  be any set not containing  $R$  and closed under the Boolean operations such that no set closed under the Boolean operations contains it but not  $R$ .

This paper is a long way from the final word on both of these topics and a complete list of outstanding issues would be rather long. But to mention just three:

1. AGM theories give axiomatic characterisations of the operations of revision, expansion and contraction and it would be useful to do the same for the operations examined here. It is obvious where to start; namely with the axioms of probability and desirability, plus the various Rigidity conditions. But as noted before, this does not uniquely determine withdrawal of opinion and no doubt there is more to be said here.

2. Not only attitudes, but conditional attitudes too can change as a result of experience. Some attention to this question is given in Bradley [4], but this needs to be extended to the treatment of non-opinionated states of mind.

3. Most importantly perhaps, there is a something of an asymmetry between the permitted indeterminacy of the agents' states of mind and the determinacy of the kinds of constraints on attitude change imposed by experience that are considered in this paper. But experience too can be indeterminate and in ways that are not simply handled by conditioning on very coarse-grained partitions. A more systematic investigation of the types of possible experience and their implications for attitude changes would be helpful.

## 7 Appendix: Proofs

Let  $A = \{A_i\}$  be any partition and let  $\alpha\beta = \{\alpha_i, \beta_i\}$  be any assignment of probability and desirability across  $A$ .

**Theorem 10** *Let  $S = \langle p, v \rangle$  be an opinionated state of mind such that the  $p(A_i) = \alpha_i$  and the  $v(A_i) = \beta_i$ . Let  $S_{\{A_i\}}^-$  be the state of mind obtained by opinion withdrawal on  $S$ . Then (a) if  $\langle p_j, v_j \rangle \in S_{\{A_i\}}^-$  then  $v_j(\cdot|A_i) = v(\cdot|A_i)$  and  $p_j(\cdot|A_i) = p(\cdot|A_i)$ ; and (b) if  $v_j(\cdot|A_i) = v(\cdot|A_i)$  and  $p_j(\cdot|A_i) = p(\cdot|A_i)$ , then  $\langle p_j, v_j \rangle \in S_{\{A_i\}}^-$ .*

**Proof.** (a) By definition, for any  $\langle p_j, v_j \rangle \in S_{\{A_i\}}^-$  there exists  $\beta = \{\beta_i\}$ , such that  $\beta_i \in \mathfrak{R}$  and  $v_j(X) = \sum_i [v(X|A_i) + \beta_i] \cdot p_j(A_i|X)$ . Hence by the definition of conditional desirability, for any  $\hat{A} \in A$ :

$$\begin{aligned} v_j(X|\hat{A}) &= v_j(X\hat{A}) - v_j(\hat{A}) \\ &= \sum_i [v(X\hat{A}|A_i) + \beta_i] \cdot p_j(A_i|X\hat{A}) - \sum_i [v(\hat{A}|A_i) + \beta_i] \cdot p_j(A_i|\hat{A}) \\ &= v(X\hat{A}|\hat{A}) + \beta_i - v(\hat{A}|\hat{A}) - \beta_i \\ &= v(X|\hat{A}) \end{aligned}$$

since  $p_i(A_i|X\hat{A})$  for all  $A_i \neq \hat{A}$ . But by Theorem 2 in Bradley [2], if  $v(\cdot|\hat{A}) = v_j(\cdot|\hat{A})$  then  $p(\cdot|\hat{A}) = p_j(\cdot|\hat{A})$ .

(b) Now suppose that  $\langle p', v' \rangle$  is such that  $v'(\cdot|A_i) = v(\cdot|A_i)$  and  $p'(\cdot|A_i) = p(\cdot|A_i)$ . Let  $\alpha_i = p'(A_i)$  and  $\beta_i = v'(A_i)$ . Then  $p'(X) = \sum_i \alpha_i \cdot p'(X|A_i)$

$= \sum_i \alpha_i \cdot p(X|A_i)$  and  $v'(X) = \sum_i [v'(X|A_i) + \beta_i] \cdot p'(A_i|X) = \sum_i [v(X|A_i) + \beta_i] \cdot p(A_i|X)$ . Hence by the definition of opinion withdrawal,  $\langle p', v' \rangle \in S_{\{A_i\}}^-$  ■

**Theorem 11** *Let  $S = \langle p, v \rangle$  be any opinionated state of mind. Then:*

$$S_{\alpha\beta}^* = (S_{\{A_i\}}^-)_{\alpha\beta}^+$$

**Proof.** If  $p(A_i) = 0$ , for any  $A_i \in A$ , then the equivalence holds trivially in virtue of the fact that  $S_{\alpha\beta}^* = \emptyset = S_{\{A_i\}}^-$ . So suppose that this is not the case. By definition  $S_{\alpha\beta}^+(S_{\{A_i\}}^-) = \{\langle p_j, v_j \rangle \in S_{\{A_i\}}^- : p_j(A_i) = \alpha_i \text{ and } v_j(A_i) = \beta_i\}$ . But by Rigidity of Withdrawal, if  $\langle p_j, v_j \rangle \in S_{\{A_i\}}^-$ , then  $v_j(X|A_i) = v(X|A_i)$  and  $p_j(\cdot|A_i) = p(\cdot|A_i)$ . Then  $S_{\alpha\beta}^+(S_{\{A_i\}}^-)$  is the pair  $\langle p_j, v_j \rangle \in S_{\{A_i\}}^-$  such that  $p_j(A_i) = \alpha_i$ ,  $v_j(A_i) = \beta_i$ ,  $v_j(X|A_i) = v(X|A_i)$  and  $p_j(\cdot|A_i) = p(\cdot|A_i)$ . Notice that the pair  $\langle p_j, v_j \rangle$  is unique since the probability and desirability of any other prospect  $X$  is determined by the probability and desirability of the  $A_i$  and the conditional probability and desirability of  $X$  given the  $A_i$ . Now  $p_{\alpha\beta}^*(A_i) = \alpha_i$  and  $v_{\alpha\beta}^*(A_i) = \beta_i$ . But since generalised conditioning satisfies Rigidity with respect to the  $A_i$ , it follows that  $p_{\alpha\beta}^*(\cdot|A_i) = p(\cdot|A_i)$  and  $v_{\alpha\beta}^*(\cdot|A_i) = v(\cdot|A_i)$ . So  $\langle p_{\alpha\beta}^*, v_{\alpha\beta}^* \rangle = \langle p_j, v_j \rangle$ . ■

**Corollary 12** *(Recovery) Suppose that  $p(A_i) = \alpha_i$  and  $v(A_i) = \beta_i$ . Then  $S = (S_{\{A_i\}}^-)_{\alpha\beta}^+$ .*

**Proof.** Follows from the fact that  $S = S_{\alpha\beta}^*$  when the new assignment of probability and desirability is the same as the old. ■

**Lemma 13** *Let  $S = \{\langle p_j, v_j \rangle\}$  be any state of mind. For any  $\langle p_j, v_j \rangle \in S$ , let  $S_j = \{\langle p_j, v_j \rangle\}$ . Then (a)  $S_{\alpha\beta}^* = \bigcup_j ((S_j)_{\alpha\beta}^*)$ , (b)  $S_{\{A_i\}}^- = \bigcup_j ((S_j)_{\{A_i\}}^-)$ , and (c)  $S_{\alpha\beta}^+ = \bigcup_j ((S_j)_{\alpha\beta}^+)$ .*

**Proof.** (a) Follows immediately from the definition of  $S_{\alpha\beta}^*$ . (b) Follows immediately from the definition of  $S_{\{A_i\}}^-$ . (c)  $\langle p_j, v_j \rangle \in S_{\alpha\beta}^+$

$$\begin{aligned} &\Leftrightarrow \langle p_j, v_j \rangle \in S, p_j(A_i) = \alpha_i \text{ and } v_j(A_i) = \beta_i \\ &\Leftrightarrow \exists S_j \subseteq S \text{ such that } S_j = \{\langle p_j, v_j \rangle\}, p_j(A_i) = \alpha_i \text{ and } v_j(A_i) = \beta_i \\ &\Leftrightarrow \exists S_j \subseteq S \text{ such that } S_j = (S_j)_{\alpha\beta}^+ \end{aligned}$$

■

**Corollary 14** *(Weak Monotonicity) Let  $S$  and  $T$  be any two states of mind. Then  $S \subseteq T \Rightarrow S_{\alpha\beta}^*/S_{\{A_i\}}^-/S_{\alpha\beta}^+ \subseteq T_{\alpha\beta}^*/T_{\{A_i\}}^-/T_{\alpha\beta}^+$*

**Theorem 15** *Let  $S = \{\langle p_j, v_j \rangle\}$  be any state of mind. Then*

$$S_{\alpha\beta}^* = (S_{\{A_i\}}^-)_{\alpha\beta}^+$$

**Proof.** For any  $\langle p_j, v_j \rangle \in S$ , let  $S_j = \{\langle p_j, v_j \rangle\}$ . Then by Lemma 13 (a),  $S_{\alpha\beta}^* = \cup_j((S_j)_{\alpha\beta}^*) = \cup_j(((S_j)_{\{A_i\}}^-)_{\alpha\beta}^+)$  by Theorem 11. But by Lemma 13 (c),

$$\begin{aligned} \cup_j((S_j)_{\{A_i\}}^-)_{\alpha\beta}^+ &= (\cup_j((S_j)_{\{A_i\}}^-))_{\alpha\beta}^+ \\ &= (S_{\{A_i\}}^-)_{\alpha\beta}^+ \end{aligned}$$

by Lemma 13 (b). ■

## References

- [1] Bolker, E. (1966) “Functions Resembling Quotients of Measures”, *Transactions of the American Mathematical Society* 124: 292-312
- [2] Bradley, R. (1999) “Conditional Desirability”, *Theory and Decision* 47: 23-55
- [3] Bradley, R. (2005) “Radical Probabilism and Mental Kinematics”, *Philosophy of Science* 72: 342-364
- [4] Bradley, R. (2007) “The Kinematics of Belief and Desire”, *Synthese* 156: 513-535
- [5] Bradley, R. (forthcoming) “Becker’s Thesis and three models of preference change”, *Politics, Philosophy and Economics*
- [6] Gärdenfors, P. (1988) *Knowledge in Flux*, Cambridge MA, MIT Press
- [7] Good, I.J. (1952) “Rational Decisions”, *Journal of the Royal Statistical Society* B 14: 107-114
- [8] Good, I.J. (1983) *Good Thinking*, Minneapolis MN, University of Minnesota Press
- [9] Hansson, S-O. (1995) “Changes in Preference”, *Theory and Decision* 38: 1-28
- [10] Hansson, S-O. 2001. *The Structure of Values and Norms*. Cambridge: Cambridge University Press
- [11] Jeffrey, R. C. (1983) *The Logic of Decision*, 2nd ed, Chicago, University of Chicago Press
- [12] Jeffrey, R. C. (1983) “Bayesianism with a Human Face”, in *Probability and the Art of Judgement*, R. C. Jeffrey ed., Cambridge, Cambridge University Press, 1992
- [13] Jeffrey, R. C., ed. (1992) *Probability and the Art of Judgement*, Cambridge, Cambridge University Press

- [14] Levi, I. (1974) “On Indeterminate Probabilities”, *Journal of Philosophy* 71: 391-418
- [15] Levi, I. (1980) *The Enterprise of Knowledge*, Cambridge MA: MIT Press
- [16] Levi, I. (1986) *Hard Choices*, Cambridge: Cambridge University Press
- [17] Levi, I. (1990) “Compromising Bayesianism: A Plea for Indeterminacy”, *Journal of Statistical Planning and Inference* 25: 347-62
- [18] Nehring, K. (2000) “A Theory of Rational Decision Under Ignorance”, *Theory and Decision* 48: 205-240.
- [19] Nehring, K (2008) “Imprecise Probabilistic Beliefs as a Context for Decision-Making under Ambiguity”, mimeo
- [20] Spohn, W. (1986) “The Representation of Popper Measures”, *Topoi* 5: 69-74
- [21] Steele, K. (2007) “Distinguishing indeterminate belief from ‘risk-averse’ preferences”, *Synthese*, 158(2):189–205, 2007.
- [22] Voorbraak, F. (2000) “Partial Probability: Theory and Applications”, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 8(3): 331-346
- [23] Voorbraak, F. (1996) “Probabilistic Belief Expansion and Conditioning”, *ILLC Research Report LP-96-07*, University of Amsterdam
- [24] Walley, P. (1991) *Statistical Reasoning with Imprecise Probabilities*, London: Chapman & Hall