Probabilism, Representation Theorems, and Whether Deliberation Crowds Out Prediction

Edward Elliott

University of Leeds

Decision-theoretic representation theorems have been developed and appealed to in the service of two important philosophical projects: (i) in attempts to characterise credences in terms of preferences, and (ii) in arguments for probabilism. Theorems developed within the formal framework that Savage (1954) developed have played an especially prominent role here. I argue that the use of these 'Savagean' theorems create significant difficulties for both projects, but particularly the latter. The origin of the problem directly relates to the question of whether we can have credences regarding acts currently under consideration and the consequences which depend on those acts; I argue that such credences are possible. Furthermore, I argue that attempts to use Jeffrey's (1990) non-Savagean theorem (and similar theorems) in the service of these two projects may not fare much better.

1. Introduction

Two questions lie at the heart of formal epistemology and contemporary decision theory, at least if we take talk of credences (or degrees of belief) seriously:

1 *What are credences?*[1]

2 *What should our credences be like?*

It would be fair to say that the following has been (and still is) a common and important answer to the first question:

Credences can be characterised largely (if not entirely) in terms of preferences

We might call this view *preference functionalism*. The sense of 'preference' involved here is usually taken to be, or to be very closely linked to, a kind of behavioural-dispositional state relating to agents' choices between possible courses of action—in Savage's words, "Loosely speaking, [to say that a person prefers α to β] means that, if he were required to decide between α and β, no other acts being available, he would decide on α" (1954, p. 17). As I understand it,

[1] More specifically, the question here concerns the conditions under which an agent counts as being in such-and-such a credence state. The orthodoxy is that belief is a binary relation between a subject at a time and a proposition, and most I imagine would be happy to extend this approach to credences: a credence of *n* towards *P* is a ternary relation between a subject at a time, a degree of confidence (represented by *n*), and a proposition *P*. The real philosophical meat lies in specifying the conditions under which an agent stands in such a relationship.

preference functionalism need not be an anti-realist, constructivist, or behaviourist position. It should be treated as neutral with respect to whether credences *are* preference states, or if credences are to be functionally characterised in significant part in terms of the preference patterns that they typically *give rise to*. The ever-present *betting interpretation* is one instance of preference functionalism—but so too would be, say, an *interpretivism* according to which the correct assignment of credences and utilities depends crucially (if not wholly) on which assignment constitutes the best overall rationalisation of the subject's behavioural dispositions (cf. Lewis 1974, pp. 337-8). Included amongst the ranks of preference functionalists are Ramsey (1931), Savage (1954), Eells (1982), Maher (1993, 1997), Davidson (2004), and Cozic and Hill (2015).

Regarding the second question, the orthodox answer is *probabilism*. In slogan form:

Credences ought to conform to the axioms of the probability calculus

More precisely, agents ought to be such that their *total credence state* can be accurately modelled by a *credence function* ($\mathcal{C}r$) which takes us from some algebraic propositional structure into the set of real numbers in the unit interval, such that for all propositions P and Q in the algebra:[2]

i) If P is epistemically necessary, then $\mathcal{C}r(P) = 1$
ii) If P is epistemically impossible, then $\mathcal{C}r(P) = 0$
iii) If $P \wedge Q$ is epistemically impossible, then $\mathcal{C}r(P \vee Q) = \mathcal{C}r(P) + \mathcal{C}r(Q)$

In all that follows, I will use the expression 'probability function' only for functions satisfying these conditions; a 'credence function' more generally being any function defined on a set of propositions intended to represent (some part of) an agent's total credence state.

Decision-theoretic representation theorems have played a prominent role in arguments for both preference functionalism and probabilism; that is, they have found within philosophy both a *descriptive* application and a *normative* application. Roughly, the typical representation theorem for *classical expected utility* (CEU) theory asserts that if a person's preferences satisfy certain constraints C, then she can be *represented* as maximizing her expected utility given a particular set of credences $\mathcal{C}r$ and utilities $\mathcal{U}$—and, moreover, that having those credences and utilities is the *only* way that she could be represented as an expected utility maximiser. (I will add clarifications to this gloss below.) These theorems come in all shapes and sizes, however,

[2] What is it for $\mathcal{C}r$ to be an *accurate* model of one's credences? I have intentionally left this matter somewhat vague to allow for variation amongst individual theorists. (Some of this variation will be discussed further in §3.) The question hinges in part upon one's metaphysics of the graded attitudes, and on one's account of just how—and to what extent—credence and utility functions are supposed to faithfully represent those attitudes.

and—as I will argue—they are not all equally well-suited for the descriptive and normative applications to which they have sometimes been put.

For the majority of this paper, I will focus on the most common kind of representation theorem: those which take preferences to be defined over 'acts', modelled as functions from states to outcomes *à la* Savage (1954). In §2, I will provide an overview of the basic formal structures that underlie any broadly 'Savagean' theorem, and provide a somewhat more precise characterisation of what a representation theorem can be taken to imply about our preferences and how they might be represented. Then, in §3, I will take a look the relevant descriptive and normative applications of these theorems.

§4 contains my main arguments against the use of Savagean theorems. In brief, such theorems do not constrain the assignment of credences to enough propositions, or to the right kinds of propositions, to adequately represent *any* realistic subject's total credence state. In the Savagean system, probabilities are only assigned to *disjunctions of states*, and many of the most interesting propositions—including our credences about acts and about outcomes—cannot be expressed as a disjunction of states.[3] This characteristic of Savage's theorem is usually attributed to his assumption that states are act-independent, in a sense to be made clear below. However, the same property attaches to all (single-primitive) theorems which make use of the same basic formal structures that Savage employs—even those which allow for act-dependent states. I will argue that this creates significant difficulties for any representation theorem argument for probabilism based on a Savagean theorem, and makes it somewhat more difficult than is ordinarily assumed to establish a plausible preference functionalist position on the back of such a theorem.

Part of my argument hangs on the issue of whether 'deliberation crowds out prediction'; i.e., on whether we can have credences regarding acts currently under our control (and regarding any propositions which are evidentially dependent on the performance of those acts). This thesis has been advocated in (Spohn 1977), (Levi 1989), (Gilboa 1994), and more recently, (Price 2012), and (Ahmed 2014). Some of those who accept the 'crowding out' thesis have even taken it as a unique *advantage* of the Savagean system that it does not involve credences for acts. In §5, I argue that the thesis is insufficiently motivated and implausible.

Finally, in §6, I will argue that things do not look much more promising if we instead appeal to a theorem like Jeffrey's (1990), which takes preference to be defined over arbitrary propositions rather than acts. Given the theorems we currently have, taking this route will only lead to problems elsewhere. As we will see, it is not so easy to do without Savage.

[3] That the probability function derived using Savage's theorem in particular does not directly supply credence values for acts has been noted before (e.g., Spohn 1977, pp. 117-8, Joyce 1999, p. 117), though the relevance of the point for the theorem's use in arguments for probabilism and as a basis for preference functionalism has not been discussed.

One point of clarification is in order before we move on: this paper is *not* intended as an argument that probabilism is false, nor that preference functionalism is implausible. I am sympathetic to both positions. Indeed, I think that representation theorems can and should be developed with specifically descriptive applications in mind, and there are many arguments in favour of probabilism which do not go *via* representation theorems.

2. The Typical Representation Theorem

The standard model of a decision situation takes the form of a matrix:

Possibilities / **Options**	*possibility* 1	*possibility* 2	*possibility* 3
option α	*outcome* 1	*outcome* 2	*outcome* 3
option β	*outcome* 4	*outcome* 5	*outcome* 6

Encoded in this model are several key elements:

* A number of possible *ways the world might be*, of which we are uncertain
* A number of *options* that we aim to choose between; typically, these are different *acts* an agent might make
* A number of *outcomes*, the consequences of a choosing a particular option given a way the world might be; these are the items we ultimately care about when choosing

The purpose of the decision matrix is to determine a *preference ranking* on the various options using some *decision rule*. According to CEU, that rule is *expected utility maximisation*: possibilities should be assigned probabilities, outcomes assigned real-valued utilities, and the preferred option should have the maximal probability-weighted average for its associated outcomes. Other *non-classical* theories of decision-making will posit distinct decision rules, and may not require credences to be probabilities or utilities to be real-valued.

Every decision-theoretic representation theorem will formalise these elements—the *possibilities*, *outcomes*, *options*, *preference ranking*, and the *decision rule*—in one way or another, though there is a significant amount of variation here. I will in this section confine myself to a discussion of Savage's CEU theorem, though the reader should keep in mind that the majority of theorems for both CEU and an enormous variety of non-classical decision theories employ the same basic formal machinery that Savage developed (see §4 for more discussion). In particular, all Savagean theorems are built around the following two sets:

* A set of *states*, $\mathcal{S} = \{s_1, s_2, \ldots\}$; i.e., a finely-grained partition of some possibility space. From $\mathcal{S}$ we construct a set of *events*, $\mathcal{E} = \{E_1, E_2, \ldots\}$, an algebra of subsets of $\mathcal{S}$ and the domain of the credence function $\mathcal{C}r$

* A set of *outcomes*, $\mathcal{O} = \{o_1, o_2, \ldots\}$. These are the objects of the utility function $\mathcal{U}$, and usually taken as descriptions of the consequences of a choice that are maximally specific with respect to what the agent cares about

$\mathcal{S}$ and $\mathcal{O}$ are then used to construct a space of options, $\mathcal{A} = \{\mathcal{F}_\alpha, \mathcal{F}_\beta, \ldots\}$. $\mathcal{A}$ consists of total functions from $\mathcal{S}$ to $\mathcal{O}$, with each $\mathcal{F}_\alpha$ in $\mathcal{A}$ intended to represent an act that the agent might perform—the supposition being that acts can be identified by their possible outcomes, conditional on the different states of the world (see Savage 1954, p. 14). More precisely, if $\mathcal{F}_\alpha$ is the function that pairs s_1 with o_1, s_2 with o_2, and so on, then it serves to represent the act α such that, were it performed, then if s_1 were the case, o_1 would result, and if s_2 were the case, o_2 would result, and so on.

The use of functions from states to outcomes to represent acts was one of Savage's most important innovations, and the characteristic feature of the formal paradigm he developed. These 'act-functions' constitute the basic relata of Savage's preference relation, $\succcurlyeq$. The notion of preference being employed here is thus intended as behavioural in character: an agent's $\succcurlyeq$ ranking is supposed to encode her choice dispositions, and so the relata of $\succcurlyeq$ are imagined as *objects of choice* rather than arbitrary propositions (as they are in Jeffrey's system; see §6).

Let $[E_1, o_1; \ldots; E_n, o_n]$ refer to the act-function $\mathcal{F}_\alpha$ such that if $s \in E_1$, $\mathcal{F}_\alpha(s) = o_1$, and if $s \in E_2$, $\mathcal{F}_\alpha(s) = o_2$, and so on, where $\{E_1, \ldots, E_n\}$ partitions $\mathcal{S}$. We can then define an *expected utility function*, *EU*, as follows:

$$EU([E_1, o_1; \ldots; E_n, o_n]) = \sum_i^n \mathcal{C}r(E_i)\cdot\mathcal{U}(o_i), \text{ where } \mathcal{C}r\colon \mathcal{E} \mapsto [0, 1] \text{ and } \mathcal{U}\colon \mathcal{O} \mapsto \mathbb{R}$$

With the basic formal elements thus specified, Savage lays out a number of conditions C_{SAV} for $\succcurlyeq$ on $\mathcal{A}$ that jointly ensure the existence of a probabilistic $\mathcal{C}r$ and a $\mathcal{U}$ that determine an *EU* which *represents* $\succcurlyeq$, in the sense that for all $\mathcal{F}_\alpha, \mathcal{F}_\beta \in \mathcal{A}$,

$$\mathcal{F}_\alpha \succcurlyeq \mathcal{F}_\beta \text{ iff } EU(\mathcal{F}_\alpha) \geq EU(\mathcal{F}_\beta)$$

Savage also proves the standard uniqueness result: $\mathcal{C}r$ is unique and $\mathcal{U}$ is unique up to positive linear transformation (i.e., up to a choice of unit and zero point). As utilities are routinely assumed to be measurable only on an interval scale, I will say that $\mathcal{U}$ is *effectively unique*—any positive linear transformation of $\mathcal{U}$ represents no more (or less) meaningful information than $\mathcal{U}$ does itself, so there is no important difference between them.

It is important to be clear on the sense in which $\mathcal{C}r$ and $\mathcal{U}$ are shown to be unique. In particular, regarding $\mathcal{C}r$, Savage's uniqueness condition only asserts that there is one function $\mathcal{C}r$ with certain properties (*viz., is a probability function on* $\mathcal{E}$) which, when combined with an appropriate $\mathcal{U}$ in the right way, will allow us to represent $\succcurlyeq$. The uniqueness result by no means establishes that the only way to represent $\succcurlyeq$ is *via* *EU*-maximisation using a probabilistic $\mathcal{C}r$ on that particular set of events $\mathcal{E}$ and bounded real-valued $\mathcal{U}$ on that particular set of outcomes $\mathcal{O}$; in fact it's well known that this is not the case. In brief, a uniqueness result for any representation theorem is only ever established relative to a choice of decision rule and certain assumptions about the functions which it takes as inputs.

Let us close this section with a brief note about interpretation. Usually, representation theorems are understood as telling us something about how *agents* can be represented *qua* decision-makers with certain kinds of attitudes. In Savage's case, the standard interpretation is:

> If S has preferences obeying constraints C_{SAV}, then S can be represented as an expected utility maximiser with credences $\mathcal{C}r$ and utilities $\mathcal{U}$, where $\mathcal{C}r$ is a probability function on a set of events $\mathcal{E}$ and $\mathcal{U}$ a real-valued function on a set of outcomes $\mathcal{O}$; furthermore, there is (effectively) only one such $\mathcal{C}r$ and $\mathcal{U}$ pair which 'fit' with this form of representation

There is an oft-neglected ambiguity in the statement of this interpretation, regarding whether $\mathcal{C}r$ and $\mathcal{U}$ are being taken as *complete* or *partial* models of the subject's range of credences and utilities respectively. On the former reading, the functions are understood as capturing *everything there is* to the subject's attitudes—so if $\mathcal{C}r$ (or $\mathcal{U}$) is not defined for some proposition P, then the agent is represented as lacking any credences (utilities) with respect to P. On the other hand, where $\mathcal{C}r$ and $\mathcal{U}$ are merely *partial* models, then if $\mathcal{C}r$ (or $\mathcal{U}$) is not defined for some proposition P, the agent is neither represented as not having nor lacking any credences (utilities) with respect to P. That is, on the latter reading, the subject is represented as having credences which *agree with* but may not be *exhaustively represented by* $\mathcal{C}r$ on $\mathcal{E}$, and likewise for $\mathcal{U}$ on $\mathcal{O}$; she may or may not have credences towards propositions outside of $\mathcal{E}$, but the representation is silent on what shape those credences might take or if those states exist.

Now, either interpretation of Savage's theorem is available, but of course it would only be useful to treat $\mathcal{C}r$ and $\mathcal{U}$ as complete models to the extent that we have good reasons to think that $\mathcal{E}$ and $\mathcal{O}$ contain all plausible objects of the subject's credences and utilities respectively. In the case of $\mathcal{O}$—which, recall, is a usually a set of propositions that are maximally specific with respect to what the agent cares about—it is immensely plausible that there are a vast number of propositions towards which any subject might have utilities that can't be found in $\mathcal{O}$; thus, any $\mathcal{U}$ on $\mathcal{O}$ is at best a partial model of an ordinary total utility state. More interesting for our purposes is whether Savage's $\mathcal{E}$ contains all possible objects of credence; in §4 and §5, I will argue that it does not. In the next section, I will say more about why this matters for representation theorem-based arguments for probabilism.

3. Preference Functionalism and Arguments for Probabilism

I'll begin with a quick look at a very standard representation theorem argument for probabilism, based on Savage's theorem or one very much like it (cf. Maher 1993). The first two premises are as follows:

P1 If S has preferences obeying constraints C_{SAV}, then S can be represented as an expected utility maximiser with credences $\mathcal{C}r$ and utilities $\mathcal{U}$, where $\mathcal{C}r$ is a probability function on a set of events, $\mathcal{E}$, and $\mathcal{U}$ a real-valued function on a set of outcomes, $\mathcal{O}$; furthermore, there is (effectively) only one such $\mathcal{C}r$ and $\mathcal{U}$ pair which 'fit' with this form of representation

P2 All ideally rational agents' preferences satisfy C_{SAV}

P1 is just the usual reading of Savage's theorem, for now still ambiguous between the *partial* and *complete* interpretations of $\mathcal{C}r$ and $\mathcal{U}$ noted just above. Together, **P1** and **P2** entail:

C1 Any ideally rational agent can be (effectively) uniquely represented as an expected utility maximiser with credences $\mathcal{C}r$ and utilities $\mathcal{U}$, where $\mathcal{C}r$ is a probability function on $\mathcal{E}$ and $\mathcal{U}$ a real-valued function on $\mathcal{O}$

As numerous authors have pointed out, **C1** does not tell us anything directly about any ideally rational agent's *actual* credences and utilities: there is a significant gap between the claim that S can be *represented* in a particular way, and that S really *is* that way.[4] To bridge that gap, advocates of the representation theorem argument will at this point typically put forward some form of preference functionalism, intended to establish a principle along something like the following lines:

P3 If S can be represented as an expected utility maximiser with credences $\mathcal{C}r$ with respect to some set of events and utilities $\mathcal{U}$ with respect to some set of outcomes, then that representation is accurate

By 'accurate', I mean whatever sense is required for the thesis of probabilism. I will say a few words about the preference functionalist motivations for **P3** (or something like it) shortly. To simplify the discussion, we'll assume that there's only ever one set of events ($\mathcal{E}$) and outcomes ($\mathcal{O}$) relative to which the relevant kind of representation exists. I doubt that this is true in general, but there are a number of complicated issues here and a full discussion would take us quite far afield without being very illuminating for the purposes at hand.

C1 plus **P3** jointly imply:

C2 If S is ideally rational, then S's credences with respect to $\mathcal{E}$ are accurately represented by a probability function $\mathcal{C}r$

[4] This point has been noted, for instance, by Maher (1993), Christensen (2001), Eriksson and Hájek (2007), and Meacham and Weisberg (2011).

Let's pause briefly to note the role that the uniqueness result is supposed to play at this stage of the argument. If there are multiple functions from $\mathcal{E}$ to [0, 1], say $\mathcal{C}r$ and $\mathcal{C}r'$, which disagree on the values they assign to particular propositions but nevertheless both fit equally well within a representation of S as an expected utility maximiser, then **P3** will end up implying that *both* $\mathcal{C}r$ and $\mathcal{C}r'$ accurately model S's credences. Whether this state of affairs is considered problematic will depend on what kind of non-uniqueness is involved and one's views about the determinacy of (representations of) our propositional attitudes.

For instance, if $\mathcal{C}r(P) \geq \mathcal{C}r(Q)$ if and only if $\mathcal{C}r'(P) \geq \mathcal{C}r'(Q)$, then you *might* think that $\mathcal{C}r$ and $\mathcal{C}r'$ are just two different ways of representing the same set of credences (cf. Zynda 2000, pp. 64-5). As noted in §2, Savage's own uniqueness result was established under the assumption that $\mathcal{C}r$ is a *probability* function. It is consistent with this that there may be many *non-probabilistic* functions on $\mathcal{E}$ which can combine with $\mathcal{U}$ to form an EU-maximising representation of $\succcurlyeq$. However, given that Savage's axioms determine a complete comparative probability ordering on $\mathcal{E}$ which will be reflected in the credence function of any EU-maximising representation, any non-probabilistic $\mathcal{C}r'$ will at least capture the same overall confidence ranking as the unique probability function $\mathcal{C}r$ that Savage proves exists, and this may well be enough for present purposes. Probabilism only requires that ideally rational agents' credences are accurately *representable* by the appropriate kind of probability functions, not that they are *uniquely* representable as such.

Where $\mathcal{C}r$ and $\mathcal{C}r'$ not only assign different *absolute* values to the propositions in $\mathcal{E}$ but also vary with respect to the overall *ordering* of those propositions—as occurs with the weaker uniqueness conditions of Jeffrey's theorem, for instance—things get a little more complicated. On the one hand, you might think that if this were the case then **P3**, as stated, leads us into contradictory claims about S's credences—that $\mathcal{C}r(P) \geq \mathcal{C}r(Q)$ and $\mathcal{C}r'(Q) > \mathcal{C}r'(P)$ cannot both be true inasmuch as $\mathcal{C}r$ and $\mathcal{C}r'$ are each supposed to be an accurate representation of a single subject's credences at a time. On the other hand, you might adopt the familiar interpretivist's line and argue that there can sometimes be more than one way to accurately represent a person's credences, even with respect to a single set of propositions $\mathcal{E}$ and even where the different representations disagree with respect to the overall confidence ranking. In any case, though, with a sufficiently strong uniqueness result we can sidestep these questions as irrelevant whenever the preference conditions of the theorem in question are satisfied. Depending on your background views, then, the *absolute* uniqueness of $\mathcal{C}r$ may not be essential to the argument—but to the extent that uniqueness *can* be established, it *does* tend make the argument a whole lot more straightforward.

Most presentations of the argument end with **C2**, as though **C2** were just a statement of probabilism. However, to get *that* result we also need the following, evidently implicit, assumption:

P4 $\mathcal{E}$ contains all and only the propositions towards which S has credences

In other words, the assumption is that $\mathcal{C}r$ is a *complete* model of a total credence state. The reason for this is clear: probabilistic coherence with respect to some set $\mathcal{E}$ is consistent with a high degree of *incoherence* overall, *if* $\mathcal{E}$ does not contain all of the propositions towards which the agent in question has credences. Thus, if **C2** is to imply probabilism, it had better be the case that $\mathcal{E}$ contains everything that S has credences towards.[5] Alongside **P4**, then, **C2** implies probabilism:

C3 If S is ideally rational, then S's *total credence state* is accurately represented by a probability function $\mathcal{C}r$ on some set of events $\mathcal{E}$

Now, there are plenty of places to take issue with the foregoing argument. I'll assume that we can take **P1** for granted. There are tricky questions concerning exactly how Savage's act-functions are to be interpreted *qua* representations of behaviours/acts, and what it is for $\succcurlyeq$ to hold between two act-functions, but that's nothing a bit of handwaving won't manage. There is also an enormous literature on whether ideally rational agents ought to satisfy C_{SAV}, but that's not a debate I'll enter into here. That leaves just two places to get off the boat: **P3** or **P4**.

The main point of this paper is the rejection of **P4**; the more common response to the representation theorem argument has been to reject **P3**—primarily by heaping scepticism on the preference functionalism that motivates it. We have already seen some reasons for concern about **P3** with respect to the uniqueness of Savage's $\mathcal{C}r$.[6] In the rest of this section, I want to

[5] A referee points out that whenever a theorem T establishes the existence of a probabilistic $\mathcal{C}r$ and a $\mathcal{U}$ relative to which $\succcurlyeq$ maximises EU, where $\mathcal{C}r$ is defined on $\mathcal{E}$ and $\mathcal{E}$ does not contain all objects of credence, then there will be a simple 'extension' of the theorem, T^*, according to which a probabilistic $\mathcal{C}r^*$ exists relative to an appropriately expanded set of propositions (call it $\mathcal{E}^*$, which we'll assume has an algebraic structure) which (a) *agrees with* $\mathcal{C}r$ on $\mathcal{E}$ and (b) combines with $\mathcal{U}$ to represent $\succcurlyeq$ in just the same way as $\mathcal{C}r$ did. We might then try to re-run the argument using T^*, and avoid the problems that I will raise later with **P4**.

There are two points to note about this. First, for the reasons to be discussed, $\succcurlyeq$ on $\mathcal{A}$ will place no interesting constraints on $\mathcal{C}r^*$ over $\mathcal{E}^* - \mathcal{E}$, so there will be many probabilistic $\mathcal{C}r^*$ satisfying (a) and (b), including ones which disagree with respect to the confidence ranking. Something would need to be said about the very significantly weakened uniqueness conditions here. Second, and more importantly, there will also be many *non-probabilistic* $\mathcal{C}r^*$ which also satisfy (a) and (b). In the context of an argument for probabilism, we cannot assume that only the probabilistic $\mathcal{C}r^*$ can form legitimate representations of ideally rational agents' credences. More generally: if we re-state the argument using T^*, then while it may be plausible that $\mathcal{C}r^*$'s domain covers all the required objects of credences, we have at best only shifted the bump under the rug.

[6] It's worth pointing out that there are *other* representation theorems with stronger uniqueness results than Savage's. For instance, Ramsey's (1931) theorem establishes not only that there's a unique probability function $\mathcal{C}r$ which figures in an expected utility representation of $\succcurlyeq$, but also that $\mathcal{C}r$ is the *only* function into $\mathbb{R}$ with this property. (The form of the representation is not identical to the EU function defined in §2, but is recognisably an expected utility formula.) See (Elliott forthcoming) for discussion and relevant proofs. While Ramsey's theorem is not ordinarily cast within the Savagean framework, it is straightforward to do so without any changes to the proof.

say just a few words about, and in support of, preference functionalism. The discussion is not intended to be exhaustive, but it will be important for §6, where I argue that Jeffrey-style theorems lack some of the features which make Savagean theorems attractive to preference functionalists.

P3 states that if a subject can be represented as an expected utility maximiser with such-and-such credences and utilities, then she actually has those credences and utilities. Probably the best-known strategy for justifying this claim comes from Maher (1993), and involves a kind of interpretivism according to which:

> … an attribution of [credences] and utilities is correct iff it is part of an overall interpretation of a person's preferences [over acts] that makes sufficiently good sense of them and better sense than any competing interpretation. (p. 9)

Here, "good sense" is cashed out partly in terms of Lewis's principle of *Rationalisation*:

> [A subject] should be represented as a rational agent; the belief and desires ascribed to him … should be such as to provide good reasons for his *behaviour* […] I would hope to spell this out in decision-theoretic terms, as follows. Take a suitable set of mutually exclusive and jointly exhaustive propositions about [the subject's] *behaviour* at any given time; of these alternatives, the one that comes true […] should be the one (or: one of the ones) with maximum expected utility according to the total system of beliefs and desires ascribed to [the subject] at that time… (1974, p. 337, emphasis added)

Like Lewis, Maher also appeals to a principle of *Charity*, but argues that if there exists a unique representation of *S* as a probabilistically coherent expected utility maximiser, then that is *the best* representation of *S*—it will be uniquely maximal according to both Charity and Rationalisation—from which **P3** follows. (Note, though, that it's no part of Maher's view that ordinary agents *ever* satisfy C_{SAV}, nor that ordinary agents should ever be represented as perfect expected utility maximisers.)

Now, the basis for Rationalisation, according to Lewis, is just folk psychology:

> Decision theory (at least, if we omit the frills) is not esoteric science … Rather, it is a systematic exposition of the consequences of certain well-chosen platitudes about belief, desire, preference, and choice. It is the very core of our common-sense theory of persons, dissected out and elegantly systematized. (1974, pp. 337-8)

This reflects a common core to all interpretivist positions: we are to understand the truth-conditions of propositional attitude attributions in terms of the general (if implicit) folk practices of interpretation—and that practice, as we know, very often involves rationalising *behaviour* in something like the above sense. It would, of course, be entirely natural to cash out the principle of Rationalisation using an expected utility representation theorem suited in particular to a behavioural construal of $\succcurlyeq$. This is exactly the kind of thing that a Savagean representation theorem helps us to do.

That is one way to develop preference functionalism, but not the only. Another approach might be *via* a more general *a posteriori* functionalism, which identifies credences and utilities through their causal-explanatory role in the production of intentional behaviour according to our best psychological theories. Given the large amount of evidence against descriptive construals of CEU, appeal to something like Savage's theorem would presumably be off the table: it seems clear that ordinary agents do not satisfy Savage's axioms, and more generally do not seem to consistently make decisions by calculating expected utilities with probabilistic credences. In a recent paper, Meacham and Weisberg (2011) make much of these two points, arguing against any kind of preference functionalism based on a CEU representation theorem. As they put it, the core problem they point to is that "the psychological picture at the heart of [such views] is false" (p. 654; see also p. 642, fn. 3).[7]

However, there is nothing tying preference functionalism in general to CEU theorems in particular. There are numerous representation theorems for a huge variety of non-classical utility theories that a preference functionalist might appeal to, such as those in (Schmeidler 1989), (Sarin and Wakker 1992), (Machina and Schmeidler 1992), (Casadesus-Masanell, Klibanoff *et al.* 2000), and many, many more. For that matter, there's no reason why a representation theorem argument for probabilism needs to be based on a CEU theorem: any theorem which is *compatible* with the existence of probabilistic $\mathcal{C}r$ will do the trick, so long as (a) the specific conditions under which $\mathcal{C}r$ is probabilistic are satisfied by ideally rational agents and (b) the relevant analogue of **P3** can be given a plausible justification. If the assumption of expected utility maximisation as a general decision rule for ideal and/or non-ideal agents is causing difficulties, there are ways to reformulate the argument without it.

Indeed, there are independent motivations for taking this route. For one thing, CEU theorems are fundamentally limited in their representation of credences, requiring as they do that $\mathcal{C}r$ is a probability function. If any agent is not probabilistically coherent—and most of us aren't—then no probability function can accurately represent her credences. On the other hand, many (though not all) non-classical utility theories allow for violations of the probability axioms. Furthermore, the preference conditions associated with non-classical utility theories tend to be weaker than those needed for CEU theorems and more descriptively realistic—in fact, they are often just weaker versions of Savage's preference conditions, designed specifically to accommodate the relevant empirical data. This is true, for instance, for axiomatisations of cu-

[7] These are hardly the only problems that Meacham and Weisberg point to, and there are more problems still in the wider literature. Again: a thorough defence of preference functionalism is not the point here, nor would such be possible in the available space.

mulative prospect theory (see Tversky and Kahneman 1992), one of the most empirically successful models of decision-making yet developed. For these reasons, representation theorems for non-classical theories should look very attractive to preference functionalists—independently of whatever role they might also play in arguments for probabilism.

4. Credences in Acts and Outcomes

We will begin by looking at how the problem arises in Savage's system in particular, before moving on to see how the problem spreads to any theorem which makes use of act-functions. The problem, of course, is that Savage's Cr is defined only for *events*. Every event is equivalent to some proposition—in particular, to some disjunction of states—but not every proposition corresponds to an event. Let us refer to propositions which do not correspond to an event as *non-event propositions*. The question is whether these non-event propositions form an important class, a set of propositions towards which we do (or can) have credences. There are two kinds of propositions to consider: those regarding what acts we might perform, and those regarding the outcomes that might result.

In proving his representation result, Savage makes use of two important assumptions. The first is that the states in $\mathcal{S}$ should be *act-independent*, where this implies (at minimum) that states should be carved up in such a way that whatever state obtains is logically independent of whatever act the agent might choose in their current decision situation. Furthermore, Savage's theorem requires very strong assumptions about what functions from $\mathcal{S}$ to $\mathcal{O}$ must be included in $\mathcal{A}$—in the simplest presentation, $\mathcal{A}$ is just the set of all such functions. This leads to a well-known problem with *imaginary acts*: on any viable conception of an outcome, at least some outcomes will be incompatible with at least some states, and so many functions from $\mathcal{S}$ to $\mathcal{O}$ cannot represent genuine acts. For instance, it is not even *metaphysically* possible that an agent could perform a "constant act" which brings about o regardless of what state obtains, if some states imply $\neg o$. There is much to be said about these two assumptions, especially the latter. But my aim here is not to raise old worries; in particular, we will sweep the imaginary acts problem under the rug and pretend it was never there. In effect, this is to assume that every act represented in $\mathcal{A}$ is compatible with every s in $\mathcal{S}$, and every s in $\mathcal{S}$ is consistent with every o in $\mathcal{O}$, for, otherwise, Savage's act-functions do not make sense *qua* representations of possible actions an agent might take.

That Cr is undefined for propositions regarding what acts we might perform follows immediately from the assumption of act-independence. Since the choice of (and performance of) any one act α implies foregoing the other options on the table, every state is consistent with the performance and non-performance of α. A visualisation will be useful. The entire rectangle represents the set of all possible worlds, partitioned into a number of states ($s_1 - s_6$), each of which contains six worlds—some where *α is performed* (represented by α), and some where *α is not performed* (~α).

s_1 α α α α ~α ~α	s_2 α α α α ~α ~α	s_3 α α α α ~α ~α
~α ~α α α s_4 α α	~α ~α α α s_5 α α	~α ~α α α s_6 α α

Every event corresponds to some collection or other of states. However, the proposition that *α is performed*—the set of all α worlds—does not correspond to any collection of states.

The same can be said for outcome-propositions. Since each outcome $o \in \mathcal{O}$ is distinct, if one outcome obtains then no other outcome does. By the same reasoning that we have just seen, then, states do not cut finely enough to make the relevant distinctions we need here. The same applies to *any* proposition I care about, the truth of which is at least partially dependent on my choices. For instance, suppose that some outcomes are *nice*, while other outcomes are *nasty*. Then, the proposition *something nice happens* is a non-event proposition: every state is consistent with nice things happening and also with nasty things happening, so there is no way to form that proposition as a disjunction of states. Or, perhaps I care about whether I get to eat tomorrow, and this is not guaranteed to occur independently of my actions. Then *s* will be compatible with both *I will eat dinner tomorrow* and *I won't eat dinner tomorrow*, so neither proposition is an event—although I certainly *do* have credences (high credences, in fact) that I will eat dinner tomorrow.

If we have credences regarding acts *or* outcomes—and we do—then $\mathcal{C}r$ cannot be an accurate and *complete* model of our credences. **P4** is false. Whatever it may or may not tell us about our credences, Savage's theorem *alone* cannot supply us with the foundations for a general, preference-based account of what it is to have the credences that we do in fact have. Something needs to be said to account for credences towards non-event propositions as well. Without this, both the argument for probabilism and any preference functionalism which underlies it are, *at best*, incomplete.

Now, I think the following would be a very healthy response to all this: so much the worse for trying to give an account of credences entirely in terms of choices between acts! No one should ever have *expected* that Savage's theorem would tell us everything there is to know about having credences—there are plenty of other things that theorists can and should appeal to in constructing their accounts of these attitudes. Most interpretivists (including Lewis and Maher) make appeal not only to Rationalisation but also to principles of Charity, which relate credences not to their role in the production of behaviour but instead to their role in response to evidence. Functionalists, likewise, should presumably make some appeal to 'input' conditions when formulating their definitions, and these are entirely neglected in any approach which only makes reference to the effects that credences have on choice behaviour. However, there's recognising that there are gaps to be filled in, and there's actually filling in those gaps. Maybe Savage's theorem (or a Savagean theorem more generally) can be used to supply a very significant *part* of a plausible story about what credences are, but we'll only be able to judge that when we have the rest of the details on the table.

Is it at least plausible that Savage's $\mathcal{C}r$ is an accurate but only *partial* model of a subject's credences (whenever the subject's preferences satisfy C_{SAV})? Of this I am more doubtful. The most common approaches to understanding our propositional attitudes are general with respect to content, and I see no reason not to think that our account of credences ought to be as well. That is, absent reasons to the contrary we should avoid a *disjunctive* account of credences, according to which we have one kind of story for what it is to have such-and-such credences for one class of propositions—they're entirely determined by preferences—and a wholly distinct and non-overlapping story for the rest. If we are going to use information which goes beyond preferences in fixing our account of credences for propositions *outside* of $\mathcal{E}$, then it seems only reasonable to expect that the very same information could be relevant for credences towards propositions *within* $\mathcal{E}$ as well. But then it seems unlikely that having such-and-such preferences should completely determine one's credences, even with respect to $\mathcal{E}$.

So much for the problem as it arises in Savage's own theorem. More than one implausible assumption went into the above argument; perhaps the right lesson to draw is that one would do best to not appeal to *Savage's* theorem as a foundation for preference functionalism or probabilism. However, the problem just outlined goes well beyond Savage's theorem. To see this, note that to raise the worry we do not need to assume that every state is compatible with every action, nor that every state is compatible with every outcome. If so much as *one* state is compatible with both P and $\neg P$, then neither proposition is an event. To return to our toy model, although only one state (s_1) is consistent with both *α is performed* and its negation, *α is performed* is not equivalent to any disjunction of the states:

s_1 α α α α ~α ~α	s_2 α α α α α α	s_3 α α α α α α
~α ~α ~α ~α s_4 ~α ~α	~α ~α ~α ~α s_5 ~α ~α	~α ~α ~α ~α s_6 ~α ~α

So, for the problem to arise, we only require that one or both of the following two conditions are satisfied:

A There are *acts* such that at least one state is consistent with the performance and the non-performance of that act.

B There are *outcomes* such that at least one state is consistent with that outcome obtaining and it not obtaining.

Neither **A** nor **B** imply that every state is consistent with every outcome, nor with every act. Their satisfaction is compatible, for example, with rejecting Savage's act-independence assumption, and with supposing that every outcome is simply an act-state conjunction (so that each outcome is consistent with exactly one state).

These are very weak conditions, and they are there for good reasons. The motivation for **B** is obvious. The point of decision theory applied to situations of uncertainty is to determine which choice to make on the basis of the different outcomes that each available act would have, given each of the different states that are consistent with what we know to be true. A quick glance at the standard decision matrix above will reveal that the framework is useless if every option has exactly the same outcome at every state as every other option. Dominance reasoning, for example, would be impossible, as no act could do better at a state than any other. Likewise, if **A** were false then there would be no sense in applying decision theory in the first place—each state would *determine* that a particular choice was made, so there would be no meaningful comparison of the outcomes of different acts at a state.

The origin of the problem, of course, is the formalisation of acts as functions from $\mathcal{S}$ to $\mathcal{O}$. For a Savagean theorem to avoid our worry, it would need to be the case that each state implies either P or $\neg P$, for any P that we take ourselves to (potentially) have credences about. At the very least, this would mean that the familiar Savage-style representation of actions as functions from $\mathcal{S}$ to $\mathcal{O}$ would be off the table. The majority of representation theorems—for both CEU and non-classical utility theories—closely follow Savage in this way of formalising the basic objects of preference. Some of these theorems may manage to avoid appealing to constant act-functions and act-independence, but all imply both **A** and **B**.

Luce and Krantz's (1971) theorem departs slightly from Savage's paradigm by representing actions as *partial* functions from $\mathcal{S}$ to $\mathcal{O}$, but while they explicitly reject Savage's act-independence assumption, their states are still consistent with multiple acts and outcomes. This is a direct consequence of their act-richness assumptions (axiom 1 of definition 1, p. 256). Other systems will occasionally take different approaches to representing the basic objects of preference. Suppes (1969) and Fishburn (1967), for example, take their objects of preference to be ordered pairs of act-functions, while Anscombe and Aumann (1963) make do with so-called *horse lotteries*, which are formally very similar to act-functions. Despite their differences, though, in these systems both **A** and **B** are required for the theorem to have a plausible decision-theoretic interpretation, and our problem arises.

5. Deliberation Does Not Crowd Out Prediction

Perhaps the issue is not as bad as I have made out—there are, after all, some who argue that we lack credences regarding whether we will perform one or another of the acts currently available to us in a given choice situation. Wolfgang Spohn (1977), for example, claims that "probably anyone will find it absurd to assume that someone has subjective probabilities for things which are under his control and which he can actualise as he pleases" (p. 115). Spohn's claim is that because it is entirely under her control whether *S* chooses to perform a given act or not, there is no sense in her being *uncertain*—or *certain*—about whether the act will be enacted; she simply lacks those credence states. Let us refer to this as the *Deliberation Crowds Out Prediction* (DCOP) thesis.

Furthermore, outcomes might be conceived of as being closely connected to acts, in such a way that if we were to lack credences in the latter then we might plausibly lack credences in the former. Indeed, Spohn (1977, p. 116) argues for precisely this. His argument presupposes that agents' credences can be represented by a probabilistically coherent credence function, $\mathcal{C}r$, such that for any pair of propositions P and Q in $\mathcal{C}r$'s domain,

$$\mathcal{C}r(P) = \mathcal{C}r(Q)\cdot\mathcal{C}r(P|Q) + (1 - \mathcal{C}r(Q))\cdot\mathcal{C}r(P|\neg Q)$$

If this were true, then if the agent had credences regarding some proposition P which probabilistically depends on her performance of an act α, she would be able to indirectly induce an unconditional probability regarding α using the above equality; hence, if she does not have credences regarding α, she cannot have credences for any such P. Of course, the generality of this argument is questionable: ordinary agents are not plausibly probabilistically coherent, and we certainly shouldn't assume that ideally rational agents are in an argument for probabilism.

In any case, though, the important point is that there may be ways to tie credences about outcomes to credences about acts in such a way that a lack of credences with respect to the latter plausibly leads to a lack of credences with respect to the former. For example, if outcomes were simply act-state conjunctions, then plausibly there should be no credences for outcomes inasmuch as there are no credences for acts. If so, then the truth of DCOP would certainly undermine the conclusions of §4. Indeed, the fact that $\mathcal{C}r$ will not represent credences about such things would be a particularly attractive *feature* of applying the Savagean framework—the relevant credence states never existed to begin with!

I do not share Spohn's sense of absurdity that is supposed to come with ascribing credences to *S* regarding acts that are presently under *S*'s complete control to perform, should she so choose. One of the strongest arguments (read: not based on the betting interpretation) for the DCOP thesis seems to be that credences regarding which action will be chosen in the present circumstances *play no role* in rational decision-making and so there is no theoretical reason to posit such states (Spohn 1977, pp. 114-5). Even supposing that this is true—it may be in *Savage*'s decision theory, but of course there are alternatives (e.g. Jeffrey); see also Rabinowicz

(2002, p. 112-4) and Joyce (2002) for a critique of this claim—it is one thing to say that credences about acts play no role (or *should* play no role) in decision-making and quite another to say that we simply *do not have* such credences.

By way of example, note that utilities for *events* also play no role in decision-making according to Savage's decision theory. By hypothesis, what event obtains is independent of the choice between acts, so any valuation of the events on the subject's behalf is irrelevant to her choice. It would be unreasonable to infer from this that we *do not have* utilities for events; at least, it certainly seems to me that I am able to judge which of two events I would prefer to be true, even if I know that this is entirely beyond my control. One of the central theoretical roles of utility assignments is to represent a subject's preferences over ways the world might be—that such states may not play a role in rational decision-making (according to Savage) does not mean that there is no reason to posit them at all. Likewise, I seem to be able to ascribe credences about my own actions to myself, even during deliberation. On the basis of past evidence, I know that when I am faced with the decision between caffeinated and non-caffeinated beverages, I tend to choose the former; were I in that situation now, I can be *confident* that I would do the same—and I should be able to represent such confidence in my credence function. I may even surprise myself with an herbal tea on occasion.

Indeed, denying the existence of these credence states comes with severe theoretical costs. For instance, the thesis is in conflict with the principle of Conditionalisation. The actions that we might make in *future* situations are not under our complete control *now*, and neither are the actions that we have *already* made. So we can have credences with respect to future and past actions. This is as it should be—in many circumstances, we ought to take credences about our past and future actions into account when making decisions. It is only credences about the actions that we might *now* perform which are ruled out by the thesis that deliberation crowds out prediction, as it is only those which are completely under our present control. But *this* certainly seems odd: I am confident now that I shall choose the caffeinated beverage when the option is available tomorrow; and tomorrow, after I have chosen that beverage, I shall be confident of having done so—but for that brief moment when I make the choice, my credences regarding that act will vanish from existence, only to reappear a moment later.[8] The same point, of course, applies to credences about propositions I care about which depend on my choice of act. Conditionalisation will not explain such changes; the conditionalisation model of rational changes in credence does not allow parts of our credence function to just disappear and then reappear later with a different associated degree of confidence. This is a heavy burden to bear for any putative argument for probabilism—to establish one pillar of orthodox Bayesian epistemology only by rejecting the other.[9]

[8] Thanks to Alan Hájek for this way of putting the point.

[9] For similar reasons, if we necessarily lack credences regarding acts (and outcomes!) then we are only a short step away from counterexamples to van Fraassen's (1995) General Reflection Principle and Lewis's (1980) Principal Principle.

There may be some sense in which 'deliberation crowds out self-prediction', but whatever that sense may be, it is *not* the sense in which we simply lack credences about acts. Rabinowicz (2002, pp. 92-3), for example, suggests that perhaps credences about acts "are available to a person in his purely cognitive or doxastic capacity, but not in his capacity of an agent or practical deliberator"; that is, while the agent *does* have credences about acts, *while* deliberating about what to do these credence states are (for whatever reason) cognitively inaccessible. This may be more plausible if we distinguish between *conscious* assignments of credence values to propositions from what we might call *standing* or *implicit* credence states. Alternatively, one might try to establish that *if* an agent S has credences regarding acts, then she *ought* not to *consider* or *use* those credences whilst deliberating on what to do—rational deliberation crowds out the consideration and/or application of certain credence states, perhaps, but not their *existence*.

6. Doing without Savage

The question we are left with is whether a representation theorem argument might be made any better with a different kind of representation theorem. I will not try to answer this question in great detail—there are too many factors to consider—but a few comments are warranted.

For most philosophers, the main alternative to Savage is Jeffrey. Jeffrey forgoes the Savagean multiset framework in favour of a monoset framework based around a single set of propositions $\mathcal{P}$—a set of subsets of some space of worlds $\mathcal{W}$—which jointly constitutes the domains of $\mathcal{Cr}$, $\mathcal{U}$, and $\succcurlyeq$. Jeffrey's theorem lacks strong uniqueness results, but this is perhaps not particularly troublesome—see, for instance, the closely related theorem in (Bradley 2007), which uses similar structures but has strongly unique $\mathcal{Cr}$ and $\mathcal{U}$ functions. There are, however, more significant issues here.

First of all, while Jeffrey's $\mathcal{P}$ allows for—in fact, *requires*—the inclusion of propositions about acts and outcomes, this is not yet enough to establish the relevant analogue of **P4** in §3's representation theorem argument:

P4′ $\mathcal{P}$ contains all and only the propositions towards which S has credences

The central worry is *not* that a set of events $\mathcal{E}$ does not contain propositions about acts and outcomes, but that the domain of $\mathcal{Cr}$ (whatever it may be) is not the set of propositions towards which our ideally rational subject has credences. Supposing just that $\mathcal{P}$ is some algebra or other on $\mathcal{W}$ will not ensure this result.

Perhaps we could let $\mathcal{P}$ be the powerset of $\mathcal{W}$, ensuring that it contains *every* proposition we could hope to find within $\mathcal{W}$.[10] If $\mathcal{W}$ can be appropriately characterised (and that's a big

[10] Jeffrey does require that $\mathcal{P}$, *minus* a special set of $\mathcal{N}$ of 'null' propositions, is atomless. This implies that $\mathcal{P} - \mathcal{N}$ cannot contain all subsets of $\mathcal{W}$ (e.g., it can't contain singleton sets), but any other proposition can be placed in $\mathcal{N}$ and assigned a credence of 0 in the final representation.

‘if’), there would then be no question that $\mathcal{P}$ contains *all* the propositions towards which any agent could have credences—but it may also end up containing too many propositions. For one thing, there are well-known cardinality issues with supposing that every set of worlds represents a possible object of belief (for discussion, see Lewis 1986, pp. 104-7). Furthermore, what reason has been given for thinking that ideally rational agents must have credences regarding every way the world might be? A better option for ensuring **P4′** would be to argue that an ideal agent has credences towards P if and only if she has some preferences regarding P (and that the set of propositions towards which she has preferences must be an algebra). This may well be true—but however we go about it, establishing **P4′** requires a further and quite significant additional step in the argument for probabilism.

The second issue involves the fit between Jeffrey’s system and the common motivations behind preference functionalism. Unlike Savage’s choice-based interpretation of $\succcurlyeq$, Jeffrey’s preferences are better understood as a kind of qualitative *mental* state: his objects of preference are not objects of choice, but objects of desire. Essentially, in Jeffrey’s system, $P \succ Q$ means that the subject finds the prospect of P being true *more desirable* than Q being true. Most interpretivists, however, focus on assigning mental states to agents on the basis of outwardly available facts about them—and in particular, their overall behavioural patterns (this is, after all, the kind of data we are supposed to use when interpreting one another). Lewis’ principle of Rationalisation, for example, is a claim about how we ought to understand and interpret a subject’s *behaviour*. The folk “platitude” is that we typically *act* in ways that tend to bring about what we desire given the way we think the world is—it is not nearly so platitudinous that we assign values to arbitrary propositions to match with the probability-weighted average utility of the different ways they might come true, as Jeffrey’s representation would have it.

More generally, many preference functionalists are attracted to representation theorems—particularly those that are suited to a choice-based interpretation of $\succcurlyeq$—because they offer the promise of *naturalising* the mind, of reducing the intentional (credences and utilities) to the non-intentional (e.g., behavioural dispositions). But where preferences are just another intentional mental state, it’s hard to see how the representation theorem helps to further this project at all—especially if we end up needing reference to degrees of belief when it comes to giving a reduction of preferences.

Finally, the very few representation theorems which have been developed using a monoset framework have focused primarily on normative applications—there is nothing like the great wealth of non-classical utility theories which have been developed using Savage’s framework. At least given the Jeffrey-style theorems we have available to us now, this rules out one of the more promising approaches for preference functionalism as discussed in §3. Until new theorems are developed, the preference functionalist who wants to use a Jeffrey-style representation theorem is stuck with CEU, and all the problems that come with it.

As things stand, then, it is not so easy to do without Savage. The formalisation of acts as functions from $\mathcal{S}$ to $\mathcal{O}$ is part of what makes the Savagean paradigm attractive: act-functions provide a *prima facie* straightforward means to connect acts to objects of uncertainty (states and events) and objects of utility (outcomes), all the while allowing theorists to characterise preferences over acts in a manner that appears to make the relation transparent to empirical observation. However, this model also comes with a cost. In order for these functions to represent acts, constraints must be placed on $\mathcal{S}$ (and hence $\mathcal{E}$) and $\mathcal{O}$—constraints which are ultimately manifest in restrictions on $\mathcal{C}r$ and $\mathcal{U}$. We can do away with those restrictions if we let $\succcurlyeq$ range over arbitrary propositions, as Jeffrey does, but doing so leaves us with significant hurdles yet to be overcome.[11]

[11] Thanks to Rachael Briggs, David Chalmers, Alan Hájek, Jessica Isserow, Leon Leontyev, Hanti Lin, Robbie Williams and several anonymous referees for helpful discussion and feedback. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 312938.

References

Ahmed, A. (2014). *Evidence, Decision, and Causality*. Cambridge University Press.

Anscombe, F. J. and Aumann, R. J. (1963). 'A Definition of Subjective Probability.' *The Annals of Mathematical Statistics* **34** (2): 199-205.

Bradley, R. (2007). 'A unified Bayesian decision theory.' *Theory and Decision* **63** (3): 233-63.

Casadesus-Masanell, R., Klibanoff, P. and Ozdenoren, E. (2000). 'Maxmin Expected Utility over Savage Acts with a Set of Priors.' *Journal of Economic Theory* **92**: 35-65.

Christensen, D. (2001). 'Preference-based arguments for probabilism.' *Philosophy of Science* **68** (3): 356-76.

Cozic, M. and Hill, B. (2015). 'Representation theorems and the semantics of decision-theoretic concepts.' *Journal of Economic Methodology* **22** (3): 292-311.

Davidson, D. (2004). 'Expressing Evaluations'. In *Problems of Rationality* Oxford University Press: 19-38.

Eells, E. (1982). *Rational Decision and Causality*. Cambridge University Press.

Elliott, E. (forthcoming). 'Ramsey without Ethical Neutrality: A New Representation Theorem.' *Mind*.

Eriksson, L. and Hájek, A. (2007). 'What are degrees of belief?' *Studia Logica* **86** (2): 183-213.

Fishburn, P. C. (1967). 'Preference-based definitions of subjective probability.' *The Annals of Mathematical Statistics* **38** (6): 1605-17.

Gilboa, I. (1994). 'Can free choice be known?'. In *The Logic of Strategy*. C. Bicchieri, R. Jeffrey and B. Skyrms, Eds. Oxford University Press.

Jeffrey, R. C. (1990). *The logic of decision*. University of Chicago Press.

Joyce, J. M. (1999). *The foundations of causal decision theory*. Cambridge University Press.

—— (2002). 'Levi on causal decision theory and the possibility of predicting one's own actions.' *Philosophical Studies* **110**: 69-102.

Levi, I. (1989). 'Rationality, Prediction, and Autonomous Choice.' *Canadian Journal of Philosophy* **19** (suppl.): 339-63.

Lewis, D. (1974). 'Radical interpretation.' *Synthese* **27** (3): 331-44.

—— (1980). 'A Subjectivist's Guide to Objective Chance'. In *Studies in Inductive Logic and Probability*. R. C. Jeffrey, Ed. University of California Press: 263-93.

—— (1986). *On the Plurality of Worlds*. Cambridge University Press.

Luce, R. D. and Krantz, D. H. (1971). 'Conditional expected utility.' *Econometrica* **39** (2): 253-71.

Machina, M. J. and Schmeidler, D. (1992). 'A More Robust Definition of Subjective Probability.' *Econometrica* **60** (4): 745-80.

Maher, P. (1993). *Betting on Theories*. Cambridge University Press.

—— (1997). 'Depragmatized Dutch Book Arguments.' *Philosophy of Science* **64** (2): 291-305.

Meacham, C. J. G. and Weisberg, J. (2011). 'Representation Theorems and the Foundations of Decision Theory.' *Australasian Journal of Philosophy* **89** (4): 641-63.

Price, H. (2012). 'Causation, Chance, and the Rational Significance of Supernatural Evidence.' *The Philosophical Review* **121** (4): 483-538.

Rabinowicz, W. (2002). 'Does Practical Deliberation Crowd Out Self-Prediction.' *Erkenntnis* **57**: 91-122.

Ramsey, F. P. (1931). 'Truth and probability'. In *The Foundations of Mathematics and Other Logical Essays*. R. B. Braithwaite, Ed. Routledge: 156-98.

Sarin, R. and Wakker, P. (1992). 'A simple axiomatization of nonadditive expected utility.' *Econometrica* **60** (6): 1255-72.

Savage, L. J. (1954). *The Foundations of Statistics*. Dover.

Schmeidler, D. (1989). 'Subjective probability and expected utility without additivity.' *Econometrica* **57** (3): 571-87.

Spohn, W. (1977). 'Where Luce and Krantz do really generalize Savage's decision model.' *Erkenntnis* **11** (1): 113-34.

Suppes, P. (1969). 'The role of subjective probability and utility in decision-making'. In *Studies in the Methodology and Foundations of Science: Selected Papers from 1951 to 1969* Springer: 87-104.

Tversky, A. and Kahneman, D. (1992). 'Advances in prospect theory: Cumulative representation of uncertainty.' *Journal of Risk and Uncertainty* **5** (4): 297-323.

van Fraassen, B. (1995). 'Belief and the Problem of Ulyssess and the Sirens.' *Philosophical Studies* **77**: 7-37.

Zynda, L. (2000). 'Representation Theorems and Realism About Degrees of Belief.' *Philosophy of Science* **67** (1): 45-69.