Why We Still Need the Logic of Decision
Author(s): James M. Joyce
Source: *Philosophy of Science*, Vol. 67, Supplement. Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part II: Symposia Papers (Sep., 2000), pp. S1-S13
Published by: The University of Chicago Press on behalf of the Philosophy of Science Association
Stable URL: http://www.jstor.org/stable/188653
Accessed: 09/09/2010 11:06

Why We Still Need the Logic of Decision

James M. Joyce†
The University of Michigan

In *The Logic of Decision* Richard Jeffrey defends a version of expected utility theory that advises agents to choose acts with an eye to securing evidence for thinking that desirable results will ensue. Proponents of "causal" decision theory have argued that Jeffrey's account is inadequate because it fails to properly discriminate the causal features of acts from their merely evidential properties. Jeffrey's approach has also been criticized on the grounds that it makes it impossible to extract a unique probability/utility representation from a sufficiently rich system of preferences (given a zero and unit for measuring utility). The existence of these problems should not blind us to the fact that Jeffrey's system has advantages that no other decision theory can match: it can be underwritten by a particularly compelling representation theorem proved by Ethan Bolker; and it has a property called *partition invariance* that every reasonable theory of rational choice must possess. I shall argue that the non-uniqueness problem can be finessed, and that it is impossible to adequately formulate causal decision theory, or any other, without using Jeffrey's theory as one's basic analysis of rational desire.

1. Introduction. Richard Jeffrey's *The Logic of Decision* (1983a) marks a watershed in our understanding of rational decision-making. It is a work that deserves to be mentioned in the same breath with Ramsey's "Truth and Probability" (1931), Savage's *Foundations of Statistics* (1972), and de Finetti's "Foresight: Its Logical Laws, Its Subjective Sources" (1964). *The Logic of Decision* has been a target of criticism in recent years, however. Current opinion has it that there are two flaws in Jeffrey's theory. First, there is the *non-uniqueness problem.* In contrast with other decision theories, it is not possible within Jeffrey's framework to secure expected utility representations for preferences that are *unique* up to the choice of a unit and a zero for measuring utility (except in special circumstances). This means that, on Jeffrey's account, the strengths of beliefs and desires are subject to a kind of indeterminacy that many find objectionable. Second,

†Send requests for reprints to the author, Department of Philosophy, The University of Michigan, 435 S. State St., Ann Arbor, MI 48109–1003.

Philosophy of Science, 67 (Proceedings) pp. S1–S13. 0031-8248/2000/67supp-0001$0.00

there is Newcomb's problem. In the first edition of *The Logic of Decision* Jeffrey advised agents to choose actions with an eye to securing *evidence* for thinking that desirable results will ensue. "Causal" decision theorists have argued that this fails to adequately discriminate the causal properties of acts from their merely evidential features, and that Jeffrey's theory should be rejected because it sometimes advises agents to choose acts that *indicate* desirable results without *causing* them.

Though serious, these problems should not blind us to the fact that Jeffrey's system has two real advantages over rival decision theories. First, it can be underwritten by Ethan Bolker's representation theorem (1966), the "gold standard" in this area as far as I am concerned. Second, its basic equation for computing expected utilities has a property called *partition invariance* that any reasonable account of rational choice should possess. I believe that these two features make Jeffrey's theory indispensable to any account of practical rationality. I am going to argue for three conclusions:

- The non-uniqueness problem is not as serious as it seems because (a) theories that purport to solve it do so only at the cost of making implausible assumptions about the structure of rational preference rankings, and (b) if we really want unique representations we can obtain them within Jeffrey's theory by imposing *independent* constraints on beliefs to go with the constraints we impose on preferences.
- Insofar as we seek theories of rational decision-making that apply to the sorts of choices that people might actually face, we must formulate our equations for expected utility in a partition independent way.
- While Jeffrey's account is not viable as a logic of *decision*, its underlying account of rational *belief* and *desire* must be incorporated into any reasonable decision theory, *including causal decision theory*.

2. Jeffrey's Theory of "Pure" Rationality. In defending these claims I am going to employ a highly simplified model of the decision-maker (that is a little richer than the one found in *The Logic of Decision*). I shall assume an agent whose desires are encoded in a *preference ranking*, a binary relation $X \geq Y$ that holds between propositions X and Y (in some Boolean algebra Ω) just in case, all things considered, she would rather learn of X's truth than learn of Y's truth. The agent's beliefs are encoded in her *confidence ranking*, a binary relation $X \mathrel{.{\geq}.} Y$ that holds between propositions (in Ω) just in case she is at least as confident in X's truth as she is in Y's truth. I am also going to assume that any question about an agent's rationality can be answered by looking at her preference and confidence rankings and at what she does. Practical rationality, in other words, is taken to be a matter of having beliefs and desires with certain properties and acting on them is a certain way. The first issue, that of deciding what sorts of preference rankings and confidence rankings an agent can rationally hold, is the province of the *theory of pure rationality*. The second issue, that of determining what actions the agent can rationally choose *given* her rational preferences and her beliefs, is the province of *the theory of choice*. A part of what I am going to argue is that, when all is said and done, we need *The Logic of Decision*, not as a logic of *decision*, but as a theory of "pure" rationality.

Here is Jeffrey's theory of pure rationality:

Jeffrey's Theory of Pure Rationality: A confidence/preference ranking pair (.≥., ≥) is rational only if there exists at least one pair of functions (prob, des), both defined on Ω, such that:

- prob is a countably additive probability that *weakly represents* .≥. in the sense that X .≥. Y only if prob(X) ≥ prob(Y)
- des is a real-valued utility function that weakly represents ≥ in the sense that $X \geq Y$ only if des(X) ≥ des (Y)
- prob and des jointly satisfy the *Jeffrey/Bolker Equation*: des(X) = Σ_ω prob(ω/X)des(ω) (were ω ranges over the atoms of Ω).

Although Jeffrey does not do things this way, it is useful to think of this theory as being divided into three distinct components: a *formal axiology*, or theory of rational preference, that tells us which preference rankings are rational, but does so without placing any constraints on confidence rankings; an *epistemology*, or theory of rational belief; that tells us which confidence rankings are rational without placing any constraints on the agent's preferences; and a *theory of coherence*, that tells us which preference rankings and confidence rankings can be rationally held *together*. When rewritten in these terms the theory can be expressed as follows:

Axiology: A preference ranking ≥ is rational only if there exists at least one probability/utility pair (prob, des) that satisfies the J/B-equation and whose utility weakly represents ≥.

Epistemology: A confidence ranking .≥. is rational only if there is at least one probability function, prob, that weakly represents it.

Coherence: If the probability in *every* (prob, des) pair whose utility weakly represents ≥ assigns a higher value to X than to Y, then every probability that represents .≥. should assign X a higher value than Y.

Axiology differs from Jeffrey's full theory of rationality in that it imposes constraints *only* on preferences. Even though the existence of a probability function is required for an agent's preferences to be counted rational, AXIOLOGY, taken by itself, says nothing about the relationship between this probability and the agent's confidence ranking. It requires only that there be a non-empty set D of (prob, des) pairs satisfying the J/B equation whose utility represents the agent's preferences. For future reference, let's call this the *representing set* for $\succeq$. EPISTEMOLOGY is an expression of the broadly Bayesian, or "probabilist," doctrine that rational belief is subject to the laws of probability. It says that there must be a non-empty set C of probability functions, the *representing set* for $.\succeq.$, all of whose members weakly represent $.\succeq.$.

COHERENCE is the weakest condition that yields Jeffrey's theory of pure rationality when combined with AXIOLOGY and EPISTEMOLOGY. COHERENCE makes it clear how, within Jeffrey's system, having certain desires can *commit* a person to holding certain beliefs. We can think of an agent's preferences as implicitly defining her *manifest confidence ranking* according to the rule that $X .\succeq._M Y$ iff $\text{prob}(X) \geq \text{prob}(Y)$ for *all* pairs (prob, des) $\in D$. When X ranked above Y in this ranking, an outside observer who knows everything there is know about the agent's preferences would be able to determine that she is more confident in X than in Y. Coherence says that an agent's confidence ranking must be an *extension* of her manifest confidence ranking.

Thanks to Ethan Bolker we know a great deal about what it takes to satisfy the strictures of Jeffrey's axiology. In his 1966, Bolker proved a representation theorem that entails that any preference ranking satisfying certain axiomatically specified constraints, the *Jeffrey/Bolker Axioms*, can be represented by the utility of at least one (prob, des) pair that satisfies the J/B-equation. I will not reproduce these axioms here (cf. Jeffrey 1978), but I do want to emphasize that they constrain preferences alone. As such, they do not impose any *direct* constraints on the agent's confidence ranking unless one conjoins them with COHERENCE.

There is a representation theorem, due to the French mathematician Villegas (1964), that goes some distance toward capturing the content of Jeffrey's epistemology. Suppose that a confidence ranking $.\succeq.$ satisfies both *de Finetti's Axioms of Comparative Probability*

- $X .\succeq. Y$ or $Y .\succeq. X$ for all $X, Y \in \Omega$
- $\mathrm{T} .\succeq. \mathrm{F}$
- If $X .\succeq. Y$ and $Y .\succeq. Z$, then $X .\succeq. Z$
- If Z is incompatible with X and Y, then $X .\succeq. Y$ iff $X \vee Z .\succeq. Y \vee Z$

as well as

Continuity: If $X_1, X_2, X_3, \ldots \in \Omega$ and if $(X_1 \vee X_2 \vee \ldots \vee X_n) .\preceq. Y$ for each n, then $(X_1 \vee X_2 \vee X_3 \vee \ldots) .\preceq. Y$

and

Non-Atomicity: If X .>. **F**, then X .>. (X & Y) .>. **F** for some $Y \in \Omega$.

Villegas proved that under these conditions there will be a *unique* probability function that represents .≥..[1]

One does not find axioms like these anywhere in *The Logic of Decision*. This is because (at the time of the first edition) Jeffrey was in the grips of a false, but widely endorsed, picture of the relationship between rational belief and desire. In the broadly pragmatist tradition of Ramsey, de Finetti, and Savage, one seeks to justify the basic tenets of probabilist epistemology by deriving them from constraints on rational desire. The goal is to show that any agent whose preferences are rational, and whose beliefs are related to her preferences in the proper way, will automatically have a confidence ranking that is probabilistically representable. In the context of Jeffrey's theory, a *pragmatic vindication of probabilism* of this type would consist in showing that EPISTEMOLOGY can be deduced from AXIOLOGY and theory of COHERENCE.

The broad structure of *The Logic of Decision* might suggest that this sort of pragmatism is what Jeffrey had in mind. After all, the only axioms to be found in the book constrain rational preference. Nevertheless, the kind of pragmatism described here does not sit comfortably within Jeffrey's system. This is where non-uniqueness becomes a worry. Most expected utility theories are founded on representation theorems that purport to deliver representations for preferences in which the probability in unique and the utility is unique up to the choice of a zero and unit. Within the Jeffrey/Bolker framework, in contrast, there is a degree of freedom in the choice of a representation that is not exhausted even when a unit and zero for utility is fixed. The key result is found in Bolker's 1966.

Bolker's Uniqueness Theorem: If the pair (prob_0, des_0) represents ≥ in the way required by AXIOLOGY, and if c is any real number for which the quantity $F(X, c) = 1 + c\,\text{des}_0(X)$ is uniformly positive. Then, the pair (prob_c, des_c) defined by

Bolker's Equations:
$$\text{prob}_c(X) = \text{prob}_0(X)F(X, c)$$
$$\text{des}_c(X) = \text{des}_0(X)[(1 + c)/F(X, c)]$$

also represents ≥ in the way required by AXIOLOGY.

1. This result is stronger than is needed for EPISTEMOLOGY. It delivers a *unique* representation while Jeffrey, a "modest" Bayesian to use Mark Kaplan's phrase (1996, 21), does not insist that beliefs be uniquely representable. To capture the "human face" of Jeffrey's Bayesianism, one must interpret these axioms as *principles of coherent extendibility*. See Jeffrey 1993b.

This theorem shows that, except in the extreme case where $\succeq$ is *unbounded* (so that every utility that represents it is unbounded), it is impossible to assign either unique subjective probabilities or utilities that are unique up the choice of scale on the basis of what the agent prefers. Since there are compelling reasons to think that agents should never have unbounded preference rankings (e.g., the utility-scaled version of the St. Petersburg paradox), non-uniqueness is going to be the order of the day in Jeffrey's theory.

To see why this undermines a pragmatic interpretation of Jeffrey's epistemology, consider the following consequence of non-uniqueness.

> **Result 1**: If $\succeq$ is *bounded*, then the pair $(.\succeq., \geq)$ can satisfy Axiology and Coherence even though $.\succeq.$ cannot be represented by *any* probability function.

I will not prove this result in detail here (cf. Joyce 1999, 134–136, for the proof), but the basic idea is straightforward. In the presence of Coherence, an agent's preferences force her to hold those comparative beliefs that appear in her *manifest* confidence ranking, but it does *not* require the converse. As long as her preferences are bounded there will be propositions that her manifest confidence ranking does not compare, so that neither $X .\succeq._M Y$ nor $Y .\succeq._M X$ holds. The agent is free to adopt any beliefs she likes regarding the relative likelihoods of these propositions, even beliefs that contravene the laws of probability (e.g., $X .\succ. Y$ and $\neg X .\succeq. \neg Y$), and she can get away with it without violating either Axiology or Coherence.

There are a number of ways to "spin" this result. I think it shows that an adequate theory of pure rationality must treat probabilism as an *independent* requirement of rationality that *cannot* be reduced to constraints on preferences. Pragmatist proponents of probabilism can try to avoid this conclusion by either (i) denying that rational agents hold beliefs that are not directly manifested in their preferences, or (ii) claiming that the inability of Jeffrey's theory to deliver unique representations constitutes sufficient grounds for rejecting it. Neither alternative is acceptable. The idea in (i) would be to supplement Coherence with its converse.

> **Manifestation**: If $\succeq$ satisfies Axiology and Coherence, then $X .\succeq. Y$ *only if* $\text{prob}(X) \geq \text{prob}(Y)$ for *all* pairs (prob, des) $\in D$.

Manifestation says that a rational agent holds *only* those opinions that are forced upon her by her desires, so that her belief state wholly *supervenes* on her preferences. This, of course, makes it impossible to satisfy Coherence and Axiology while holding beliefs that violate the laws of probability.

Manifestation should be rejected. The idea that preferences determine beliefs is a holdover from the bad-old behaviorist days when it was thought that (a) psychology should deal solely in observable magnitudes, (b) it is possible to analyze preferences reductively in terms of behavior, and (c) one can understand everything there is to understand about the rationality of beliefs by considering their role in the production of action. I think all three of these claims are false, and it has always amazed me that Ramsey, de Finetti, Savage, and other early proponents of expected utility theory were able to arrive at such a powerful and compelling account of rational thought and action starting from such misguided methodological assumptions. I will not, however, argue against behaviorism here. For present purposes, the important point is that it is not really an option for anyone who endorses Jeffrey's theory of rational preference.

When we embrace MANIFESTATION in the context of Jeffrey's theory we are forced to say that a person with bounded preferences *can*not adopt views about the comparative likelihoods of propositions at opposite ends of her preference ranking. If only she felt things more strongly, so that the relative difference between the very desirable and the very undesirable were greater for her, MANIFESTATION would allow her to adopt such views, but the restrained character of her wants restricts her ability to believe. This cannot be! One's ability to hold beliefs is never a function of how strongly one prefers good things to bad.[2] A Buddhist who has succeeded in extinguishing all her desire can still regard certain things as evidence for other things, think that certain events are unlikely, and so on, all without having a single desire. MANIFESTATION is in no way out for those who endorse Jeffrey's account of rational preference.

Perhaps Jeffrey's theory is the problem. Ultimately, it is the non-uniqueness of Jeffrey's representations that causes all the trouble; the proof of Result 1 does not go through when only a single probability is consistent with the agent's preferences. Since other theories deliver *unique* representations, thereby letting us deduce the basic probabilist constraint on rational belief from constraints on rational preference, pragmatist Bayesians may choose to reject Jeffrey's account in favor of some theory (Ramsey 1931, Savage 1972, Fishburn 1973, Anscombe and Aumann 1963, Luce and Krantz 1971) that can deliver unique representations. This is a bad idea for two reasons. First, when one looks closely at the way in which these theories obtain unique representations what one finds is mostly smoke and mirrors. This is a large issue, which I have discussed in detail elsewhere (Joyce 1999, §3.3), but the basic point is that unique representations are secured only by making highly implausible assumptions about the complexity of the set of prospects over which the agent's preferences are defined. For any two events X and Y that figure in the agent's

2. According to Jeffrey, the first person to make this point was Isaac Levi.

confidence ranking, there must be prospects A and B such that A's outcome when X obtains (fails to obtain) is exactly as desirable as B's outcome when Y obtains (fails to obtain).[3] This sometimes occurs, but it must happen in *every single case* if we are to extract a unique probability representation from a set of preferences. When X and Y sit at opposite ends of a preference ranking, however, there will often be no remotely plausible way to find suitable prospects to serve as A and B. The moral is that the non-uniqueness of probability representations in Jeffrey's theory provides us with no reason to reject it in favor of some other theory since no other theory really does any better.

Even if this were not so, there would still be a second, stronger reason not to give up on Jeffrey's theory. Jeffrey's rule for assigning utilities is *partition invariant* in the sense that a prospect's desirability can always be written as $\text{des}(X) = \Sigma_i \text{prob}(X/E_i)\text{des}(X \,\&\, E_i)$ where $\{E_1, E_2, E_3, \ldots\}$ can be *any* countable set of mutually exclusive, jointly exhaustive propositions. Not all expected utility theories are expressed in a partition invariant form. Savage's theory, for example, is *not* partition invariant. His equation $\text{U}(A) = \Sigma_S \text{prob}(S)\text{u}(A(S))$ for computing expected utilities is sure to yield correct expected utilities only when it is applied to "grand world" decisions whose outcomes are individuated so finely that *everything* the agent cares about is resolved by the state of the world once she chooses an action. Partition invariance, in my view, is not an optional virtue for an expected utility theory to have, since only a partition invariant theory can be applied to the kinds of "small world" decisions that people face in real life (cf. Joyce 1999, §4.1). This leaves us with only one reasonable "spin" left to put on Result 1. Given that we want a theory of "pure" rationality that both incorporates a partition invariant expected utility and a probabilist epistemology we must be willing to augment the Jeffrey/Bolker axioms for rational preferences with constraints, like de Finetti's axioms and the Villegas principle, that apply directly to rational beliefs.

Conclusion 1: Result 1 shows that AXIOLOGY and COHERENCE do not suffice as the foundation of a theory of pure rationality. EPISTEMOLOGY must be introduced as an *independent* constraint on confidence rankings.

This conclusion presents us with a piece of unfinished business. We know what it takes to satisfy AXIOLOGY: one's preferences must obey the

3. Savage accomplishes this using "constant" acts that generate an equally desirable outcome in all circumstances, other theories rely on equally objectionable devices. In this group I include Richard Bradley's (1999) formulation of Jeffrey's theory in terms of conditionals, which relies on the problematic assumption that for any propositions X and Y there is a Z such that the agent is indifferent between X and $Y \,\&\, Z$.

Jeffrey/Bolker axioms. We know what it takes to satisfy EPISTEMOLOGY: one's confidence ranking must conform to de Finetti's axioms, Villegas' continuity axiom, and be extendible to a complete, atomless ranking. What we do not yet know is what it takes for there to be a single *joint* representation for an agent's preferences and comparative beliefs. The answer is giving by the following result (see Joyce 1999, 139).

Result 2: If $\geq$ satisfies the Jeffrey/Bolker axioms, $.\geq.$ satisfies de Finetti's axioms of comparative probability and the Villegas continuity axiom, and the two jointly satisfy

Nullity*: $X\,.=.\,(X\,\&\,-X)$ iff there is some $Y \in \Omega$ such that Y is ranked with $(Y \vee X)$ by the agent's preferences even though Y is not ranked with X.

Impartiality*. If X, Y, $Z \in \Omega$ and Z is incompatible with both X and Y, then $X\,.>.\,Y$ if either of the following patterns of preference obtain
$Z > (Y \vee Z) > (X \vee Z) > X \geq Y$, or
$Y \geq X > (X \vee Z) > (Y \vee Z) > Z$,

then there exists a *unique* countably additive probability function, prob, and a real-valued utility function, des, that jointly satisfy the J/B equation such that prob weakly represents $.\geq.$ and des weakly represents $\geq$.

Again, the proof of this theorem lies beyond the scope of this paper, but the crucial step is provided by the following result.

Lemma: Suppose that $.\geq.$ and $\geq$ satisfy Nullity* and Impartiality*. Let prob be any probability that represents $.\geq.$, and (prob_0, des_0) be any probability/utility pair obeying the J/B-equation whose utility represents $\geq$. If one defines a constant c by setting

$$c = [\text{prob}(X_0)\text{-prob}_0(X_0)]/[\text{prob}_0(X_0)\text{des}_0(X_0)]$$

where X_0 is any element of Ω for which $\text{prob}_0(X_0)\text{des}_0(X_0) > 0$, then the quantity $\text{F}(X, c) = 1 + c\text{des}_0(X)$ is uniformly positive on Ω (which means prob_c represents $.\geq.$, and $\text{prob} = \text{prob}_c$).

Result 2 shows that the non-uniqueness inherent in Jeffrey's theory of preference can be removed by imposing constraints directly on confidence rankings. Non-uniqueness is thus not an *intrinsic* feature of preferences in Jeffrey's theory; rather, it is a consequence of trying to make a theory of rational preference do something that can only be accomplished when such a theory is combined with an epistemology.

3. The Theory of Rational Choice. Let's now turn from Jeffrey's theory of pure rationality to his account of rational choice. A theory of rational choice specifies what actions an agent can rationally perform given her preferences and beliefs. In the first edition of *The Logic of Decision*, Jeffrey defended the following "evidential" principle of rational decision making.

> **Jeffrey's "Evidential" Decision Rule**: A rational decision-maker will always choose an act A that maximizes Jeffrey/Bolker utility, so that des(A) $\geq$. des(B) for all alternatives acts B.

As is well known, this rule breaks down in cases, like Newcomb's problem or Prisoner's Dilemma With a Twin, in which the act that is the best *indicator* of desirable results is not most efficacious in *causing* these results. According to proponents of *causal* decision theory, the lesson is that Jeffrey's decision rule must be rejected in favor of a rule that explicitly incorporates the agent's beliefs about what her acts are likely to *causally promote*.

Causal decision theory can be formulated in a variety of ways, which are regarded as more-or-less equivalent by its defenders (cf. Harper and Skyrms 1988, x). Here I will focus on formulation found in Gibbard and Harper 1978. On this account, rational agents must maximize "causal" expected utility defined as $U(A) = \Sigma_{\omega}\, \text{prob}(A \,\Box\!\!\rightarrow \omega)\text{des}(\omega)$ where $A \,\Box\!\!\rightarrow \omega$ is the subjunctive conditional "If A were the case, then ω would also be the case." Choosing acts according to this rule gets the right answers in Newcomb-like problems, but it also introduces irreducibly causal notions into the theory of rational choice. This is something Jeffrey has always resisted. While he recognized that decision theory needs to take an agent's causal beliefs into account, he also thought it should be able to analyze such beliefs in terms of ordinary conditional probabilities. Jeffrey has made a number of serious attempts to carry out this reductive program. There are many interesting issues to be considered here, but I will not pursue them. Let me instead simply state my view without argument. I think that any reasonable decision rule must take the actor's beliefs about what her acts are likely to cause into account. Moreover, these "causal" beliefs cannot be analyzed in terms of conditional probabilities regarding propositions that lack causal content.

If you are with me on this, then you are probably wondering why it was worth making all the fuss over Jeffrey's theory. Why should a causal decision theorist worry about the formulation of a rival decision theory? The reason is that we causal decision theorists need Jeffrey's theory to properly formulate our own. The thing that makes Jeffrey's theory so indispensable to us is its partition invariance. As usually formulated causal decision theory is *not* partition invariant, which suggests that it can only be applied to decision problems that are posed in just the right way. While some may be willing to live with this (cf., e.g., Lewis 1981, 11; Sobel 1994, 161), we can do better once we recognize that Jeffrey discovered the correct theory of "pure" rationality but applied it incorrectly to the problem of choosing actions.

To find a partition invariant expression for causal decision theory one needs a way of assigning utilities to the act/state conjunctions that appear in the sorts of decision problems people actually encounter; one needs a rule that associates a causal utility $U(A \,\&\, E)$ with each element E of an event partition so that $U(A) = \Sigma_E \text{prob}(A \,\Box\!\!\rightarrow E)U(A \,\&\, E)$. The only way to do this is by setting $U(A \,\&\, E) = \Sigma_\omega \text{prob}(A \,\Box\!\!\rightarrow \omega / A \,\Box\!\!\rightarrow E)\text{des}(\omega)$. If we use the expression prob^A to denote what David Lewis calls the agent subjective probability *imaged* on A, which encodes the agent's degrees of beliefs under the *subjunctive* supposition that A will be performed, then we can rewrite this as $U(A \,\&\, E) = \Sigma_\omega \text{prob}^A(\omega/E)\text{des}(\omega)$.[4] Notice that our definition of $U(A \,\&\, E)$ is now just the Jeffrey/Bolker equation with the agent's unconditional probabilities replaced by her probabilities imaged on A. In addition to securing partition invariance, this way of expressing causal decision theory can help us appreciate the relationship between its decision rule and Jeffrey's account of pure rationality.

Causal decision theorists agree with Jeffrey about one key point concerning the nature of decision-making. As Jeffrey was first to explicitly recognize, a rational agent must evaluate each of her actions from an epistemic perspective that takes its performance into account. Jeffrey's insight, in other worlds, was that a decision-maker should seek to maximize not the *unconditional* utilities of her actions, but their utilities *conditional on the supposition that they are performed.* Here is the general principle.

> **Decision Rule**: The choice of act A is rational only if $\text{des}(A \| A) \geq \text{des}(B \| B)$ for all alternatives B, where $\text{des}(X \| C)$ is the desirability of *X conditional on C.*

Despite appearances, there is nothing here that a causal decision theorists should dispute. The dispute between causal and evidential decision theorists does not concern Jeffrey's thesis that acts should maximize conditional expected utility, but with the *sort* of conditional expected utility that they are supposed to maximize.

To explain this I need to introduce two new notions. The first is that of a generalized conditional probability function or a *supposition function*

4. Lewis 1976. Imaging is not standard conditionalization. Rather than being governed by the agent's unconditional beliefs about conjunctions involving A, as $\text{prob}(X/A)$ is, $\text{prob}^A(X)$ is related to the agent's unconditional degrees of belief about *subjunctive conditionals* according to the rule: $\text{prob}(\neg(A \,\Box\!\!\rightarrow \neg X)) \geq \text{prob}^A(X) \geq \text{prob}(A \,\Box\!\!\rightarrow X)$. One has $\text{prob}^A(X) = \text{prob}(A \,\Box\!\!\rightarrow X)$ when the law of *conditional excluded middle* holds. For more details on imaging and its role in causal decision theory see my 1999, Ch. 6.

for a subjective probability prob. This is a mapping prob(•||•) defined on $\boldsymbol{\Omega} \times \boldsymbol{C}$, where $\boldsymbol{C}$ is a distinguished subset of *conditions* in $\boldsymbol{\Omega}$, that satisfies the following for any $C \in \boldsymbol{C}$

SUP$_1$ (*Coherence*): prob(•||C) is a probability on $\boldsymbol{\Omega}$.
SUP$_2$ (*Certainty*): prob(C||C) = 1.
SUP$_3$ (*Regularity*): prob(X||C) ≥ P(C & X).

Suppositions functions describe changes in belief that occur when an agent supposes that a given proposition is true. Different functions embody different *epistemic perspectives* from which these evaluations can be made. *Bayesian* conditionalization is the only supposition function that obeys Bayes's Rule: prob(X||C)prob(Y & C) = prob(Y||C)prob(X & C). Imaging too is a supposition function, but not a Bayesian one.

Once we have a supposition rule in place we can define a notion of conditional utility, or *utility under a supposition.*

Conditional Utility: Given a supposition function prob(•||•) and an assignment of utilities to atoms of Ω, the *utility of X on the supposition that C* is defined by the *Generalized J/B-Equation*

$$\text{des}(X\|C) = \Sigma_\omega\,[\text{prob}(\omega\ \&\ X\|C)/\text{prob}(X\|C)]\ \text{des}(\omega).$$

The Generalized J/B-Equation is just the ordinary J/B-equation with the agent's unconditional probability replaced by her probability under the supposition that C obtains. I maintain that any reasonable theory of conditional utility must assume this form because any such theory must be partition invariant, and any partition invariant utility theory must have this form. Thus, a theory of conditional utility must be based on Jeffrey's theory of pure rationality in the sense that, for each condition C, a rational agent's preferences given C and beliefs given C must satisfy Axiology, Epistemology, and Coherence.

When we interpret prob(•||•) as Bayesian conditioning the Generalized Jeffrey/Bolker equation defines an "evidentialist" notion of conditional utility,

Evidential CEU: $\text{des}(X/C) = \Sigma_\omega\,[\text{prob}(\omega\ \&\ X/C)/\text{prob}(X/C)]\ \text{des}(\omega)$

and the Decision Rule tells us to maximize evidential expected utility since des(A/A) = des(A). On the other hand, if we interpret supposition as imaging we get a causalist notion of conditional utility,

Causal CEU: $\text{des}(X\backslash C) = \Sigma_\omega\,[\text{prob}^C(\omega\ \&\ X)/\text{prob}^C(X)]\ \text{des}(\omega)$

and the associated decision rule is that of causal decision theory since des(A\A) = U(A).

The fact that evidential and causal decision theory can both be written as instances of the Generalized J/B-equation shows us that the two theories *agree* about the nature of pure rationality. Jeffrey taught us that unconditional beliefs and desires must satisfy the principles of AXIOLOGY, PROBABILISM, and COHERENCE. Once we recognize that this also goes for beliefs and desires under a supposition, we come to appreciate that *all value is news value viewed from some epistemic perspective.* Though causal and evidential decision theorists may disagree about the correct epistemic perspective to adopt when evaluating actions, that is the sole source of their disagreement. Everything else we need to know about the nature of rationality was covered in *The Logic of Decision.*

REFERENCES

Anscombe, F. and R. Aumann (1963), "A Definition of Subjective Probability", *Annals of Mathematical Statistics* 34: 199–205.

Armendt, B. (1986), "A Foundation for Causal Decision Theory", *Topoi* 5: 3–19.

Bolker, E. (1966), "Functions Resembling Quotients of Measures", *Transactions of the American Mathematical Society* 124: 293–312.

———. (1967), "A Simultaneous Axiomatization of Utility and Subjective Probability", *Philosophy of Science* 34: 333–340.

Bradley, R. (1999) "A Representation Theorem for Decision Theory with Conditionals", *Synthese* 116, 2.

De Finetti, B. (1964), "Foresight: Its Logical Laws, Its Subjective Sources", in H. Kyburg and H. Smokler (eds.), *Studies in Subjective Probability.* New York: John Wiley & Sons, 93–158.

Fishburn, P. (1973), "A Mixture-Set Axiomatization of Conditional Subjective Expected Utility", *Econometrica* 41: 1–25.

Gibbard, A. and W. Harper (1978), "Counterfactuals and Two Kinds of Expected Utility", in C. Hooker, J. Leach, and E. McClennen (eds.), *Foundations and Applications of Decision Theory*, vol 1. Dordrecht: Reidel, 125–162.

Harper, W. and B. Skyrms (eds.) (1988), *Causation in Decision, Belief Change, and Statistics, II.* New York: Kluwer Academic Press.

Jeffrey, R. (1978), "Axiomatizing The Logic of Decision", in C. Hooker, J. Leach, and E. McClennen (eds.), *Foundations and Applications of Decision Theory*, vol. 1. Dordrecht: Reidel, 227–231.

———. (1983), *The Logic of Decision*, 2nd ed. Chicago: University of Chicago Press.

———. (1983), "Bayesianism with a Human Face", in J. Earman (ed.), *Testing Scientific Theories.* Midwest Studies in the Philosophy of Science, vol. X. Minneapolis: University of Minnesota Press, 133–156.

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

Kaplan, M. (1996), *Decision Theory as Philosophy.* Cambridge: Cambridge University Press.

Lewis, D. (1976), "Probabilities of Conditionals and Conditional Probabilities", *The Philosophical Review* 85: 297–315.

———. (1981), "Causal Decision Theory", *Australasian Journal of Philosophy* 59: 5–30.

———. (1986), "Probabilities of Conditionals and Conditional Probabilities II", *The Philosophical Review* 95: 581–589.

Luce, R. D. and D. Krantz (1971), "Conditional Expected Utility", *Econometrica* 39: 253–271.

Ramsey, F. (1931), "Truth and Probability", in *The Foundations of Mathematics and Other Logical Essays.* Edited by R. Braithwaite. London: Kegan Paul, 156–198.

Savage, L. (1972), *The Foundations of Statistics*, 2nd ed. New York: Dover.

Skyrms, B. (1984), *Pragmatics and Empiricism.* New Haven: Yale University Press.

Sobel, J. H, (1994), *Taking Chances: Essays on Rational Choice.* New York: Cambridge University Press.

Villegas, C. (1964), "On Qualitative Probability σ-Algebras", *Annals of Mathematical Statistics* 35: 1787–1796.