Erschienen in: Homo Oeconomicus ; 34 (2017), 4. - S. 293-311
https://dx.doi.org/10.1007/s41412-017-0060-5

Knightian Uncertainty Meets Ranking Theory

Wolfgang Spohn[1]

Abstract Knightian uncertainty is not a special kind of uncertainty; it's just uncertainty. And it raises the issue how we may model uncertainty. The paper gives a brief overview over non-probabilistic measures of uncertainty, starting with Shackle's functions of potential surprise and mentioning non-additive probabilities, Dempster–Shafer belief functions, etc. It arrives at an explanation of ranking theory as a further uncertainty model and emphasizes its additional epistemological virtues, which consist in a representation of belief, i.e., of taking something to be true (which is the basic notion of traditional epistemology and admits of degrees as well) and a full dynamic account of those degrees. The final section addresses the issue how these uncertainty measures and in particular ranking theory may be used within a decision theoretic context.

Keywords Knightian uncertainty · Functions of potential surprise · Baconian probability · Non-additive probability · Ranking theory · Belief · Decision theory

JEL Classification B21 · B50 · D81 · D83

1 Knightian Uncertainty

This paper intends to display the affinities and dissimilarities between accounts of Knightian uncertainty and ranking theory. The upshot will be that ranking theory, to be introduced below, is Janus-faced. It may be taken as a possible model of

✉ Wolfgang Spohn
wolfgang.spohn@uni-konstanz.de

[1] Fachbereich Philosophie, Universität Konstanz, 78457 Constance, Germany

Knightian uncertainty. In my preferred interpretation, though, it goes beyond Knightian uncertainty in talking rather about belief, which is usually uncertain as well, but is belief nevertheless.

We are well advised to start with clarifying the notion of Knightian uncertainty. Current economists appear to attach a specific meaning to it, which is not quite easy to state, just as it is not so easy to square economic epistemology with philosophical epistemology. Is Knightian uncertainty a special kind of uncertainty? Is it radical? I don't see this. We are more or less uncertain everywhere, even if we believe to know something. Knightian uncertainty *is* uncertainty—full stop.

Knight (1921) explains it in his crucial chapter VII (while chapter VIII discusses the individual behavioral consequences and the later chapters deal with the economic and social extensions). Chapter VII is a dense essay in philosophical epistemology and in particular in the philosophy of probability, sometimes difficult to understand, sometimes questionable, but definitely still worth reading. His understanding of risk and uncertainty is well summarized at the end (pp. 119f.).

> The liability of opinion or estimate to error must be radically distinguished from probability or chance of either type, for there is no possibility of forming *in any way* groups of instances of sufficient homogeneity to make possible a quantitative determination of true probability. Business decisions, for example, deal with situations which are far too unique, generally speaking, for any sort of statistical tabulation to have any value for guidance. The conception of an objectively measurable probability or chance is simply inapplicable. The confusion arises from the fact that we do estimate the value or validity or dependability of our opinions and estimates, and such an estimate has the same *form* as a probability judgment; it is a ratio, expressed by a proper fraction. But in fact it appears to be meaningless and fatally misleading to speak of the probability, in an objective sense, that a judgment is correct. As there is little hope of breaking away from well-established linguistic usage, even when vicious, we propose to call the value of estimates a third type of probability judgment, insisting on its differences from the other types rather than its similarity to them.
>
> It is this third type of probability or uncertainty which has been neglected in economic theory, and which we propose to put in its rightful place.

The first kind of probabilities is what he calls a priori, the second kind is called empirical. Both kinds are chances or *objective* probabilities. However, Knight correctly observes that there are many fields of reality where objective probabilities cannot be known or even do not make sense at all. These are the fields of uncertainty. So, roughly, the opposition between risk and uncertainty is the opposition between knowledge of objective probabilities on the one hand (Knight admits that we often have only estimates of objective probabilities) and subjective uncertainty on the other (which he calls a 'third type of probability judgment').

Contemporaneously, Keynes (1921) was remarkably more optimistic by developing a theory of *logical* probabilities, which he conceived as subjective probabilities objectively constrained by principles of rationality. He granted, though, that these principles are incomplete and often do not allow a comparison of probabilities. Deservedly or undeservedly, the program of logical probabilities has declined after Carnap's heroic efforts of explication, the latest version of which is found in Carnap (1971).

However, is Knight talking about probabilities in the mathematical sense, the additive structure of which was long assumed and later on succinctly axiomatized by Kolmogorov? I confess, the older the texts I read, the less sure I am whether they take "probability" in the precise mathematical or in a less determinate intuitive sense. I think Knight wrote late enough to intend that mathematical sense. Hence I tend to interpret him as saying that his uncertainty falls under these axioms as well. If so, it would seem correct to subsume Savage (1954) under the Knightian tradition, even though Savage does not refer to Knight.

The issue is not very clear, though. Risk, i.e., the kind of uncertainty that follows from knowledge of objective probabilities, must conform to these axioms, because objective probabilities are governed by them. That's clear from Knight's chapter VII (and generalizes to the case where we only have estimates of objective probabilities). Lewis (1980), by the way, reverses the argument, arguing that chances must follow those axioms, because credences must rationally do so and because credences and chances are related by his Principal Principle. However, I find no argument in that chapter VII that uncertainty must conform to those axioms as well. The quote above expresses some uneasiness with calling it probability at all. I guess he called it probability simply because he had no alternative model of uncertainty.

The prevailing interpretation of economists more determinately tends to the opposite side. It is well expressed, e.g., by Ellsberg (1961, p. 643):

> There has always been a good deal of skepticism about the behavioral significance of Frank Knight's distinction between "measurable uncertainty" or "risk," which may be represented by numerical probabilities, and "unmeasurable uncertainty" which cannot. Knight maintained that the latter "uncertainty" prevailed – and hence that numerical probabilities were inapplicable – in situations when the decision-maker was ignorant of statistical frequencies of events relevant to his decision; or when a priori calculations were impossible; or when the relevant events were in some sense unique; or when an important, once-and-for-all decision was concerned.

However, on p. 653 Ellsberg observes that Knight had considered the very same example on which Ellsberg builds his case, and he criticizes Knight for having probabilistic intuitions about it, as axiomatized later by Savage. This would rather support my interpretation of Knightian uncertainty as subjective probability.

Obviously, uncertainty is quite an indeterminate notion; it may just be an ill-defined subjective magnitude, i.e., a subjective state that somehow allows of degrees. So, the correct conclusion is perhaps that, while risk must be construed probabilistically, uncertainty need not be so construed. The issue is reflected in the decision-theoretic triad of decisions under certainty, under risk, and under uncertainty, where in the case of uncertainty the ambition was to find decision rules with no reference to epistemic states at all. This certainly threw out the baby with the bath water. Leaving uncertainty unmodeled will not do.

Hence, the serious issue—and the one I am concerned with here—is whether uncertainty, unlike risk, has other than probabilistic models (of an empirical-psychological kind or even on a rational basis). In particular, these models should also account for the dynamics of uncertainty. In the meantime, a bewildering number of such alternatives are on the table. I think the current increased interest in Knight (1921)—his Google Scholar figures have increased dramatically since 2002—is due to the fact that he first raised the issue of uncertainty and its bearing for macroeconomics, even though he was still stuck in the probabilistic paradigm.

I can't comment on the macroeconomic relevance. Here I am just interested in those alternative models of uncertainty. In the infancy of probability theory the mathematical structure was not yet fully fixed; in the 18th century Jacob Bernoulli and Johann Heinrich Lambert apparently speculated about non-additive probability (cf. Shafer 1978). This seemed soon forgotten. Surprisingly, the search for alternatives started again within economics with the so-called functions of potential surprise developed by Shackle (1949). His work was obviously unrelated to that of Knight [Shackle (1949) refers to Knight only once in a footnote and Shackle (1961) does not at all]. Still, followers of Knight like James Buchanan became very fond of Shackle's ideas. The functions were later introduced into philosophy, but used for quite different purposes by Levi (1967).

However, the field really started flourishing only in the 70s, also under the influence of AI, which emphasized the need of computable uncertainty measures; probability theory was tamed for algorithmic use only by the theory of Bayes nets (cf. Pearl 1988). So, we find Dempster–Shafer belief functions (Shafer 1976), imprecise or lower and upper probabilities, fuzzy logic and possibility theory (Zadeh 1978, Dubois and Prade 1988), also ranking theory (Spohn 1988), and more. Economists later joined again with the account of non-additive probability by Schmeidler (1989). This development is beautifully summarized, systematized, and subsumed under the most general theory of 'plausibility measures' in Halpern (2003). All this may be seen as attempts at accounting for 'Knightian uncertainty'.

I just listed ranking theory as well. In this paper I would like to discuss whether I have done so properly, that is, whether ranking theory can be interpreted as a model of Knightian uncertainty. For this purpose I first discuss Shackle's functions of potential surprise in Sect. 2, which look very much like ranking functions. Section 3 introduces ranking theory and argues that, despite its similarity, ranking theory should rather be dissociated from Shackle's and the other accounts mentioned and be viewed as pursuing other epistemological goals, which are not properly attended to in economics till the present day. So far this is pure epistemology. Hence, Sect. 4 briefly discusses the prospect of a decision-theoretic expansion of ranking theory.

2 Functions of Potential Surprise

Shackle (1961) distinguishes between ‘distributional uncertainty variables’ governed by probabilities and ‘non-distributional uncertainty variables’ governed by something else. Thus he seems to assume that reality falls into two parts, represented by two kinds of variables, for which two different kinds of epistemic attitudes are appropriate. I understand Knight as also suggesting such a division of reality. What’s the other epistemic attitude dealing with non-distributional certainty? Shackle attempts to represent it by his functions of potential surprise axiomatized in Shackle (1949, pp. 131f.) and more explicitly in Shackle (1961, p. 80). In modern language they may be described thus:

Let W be a set of mutually disjoint and jointly exhaustive possibilities or ‘possible worlds’, and let subsets of W be propositions (= Shackle’s ‘hypotheses’) (we need not consider more sophisticated algebras of measurable propositions). Then y is a *function of potential surprise* iff for all disjoint propositions A and B:

(1) $y(W) = 0$ and $y(\emptyset) = 1$,
(2) $y(A \cup B) = \min(y(A), y(B))$ (*the law of disjunction*).

Clearly, the law of disjunction extends to non-disjoint propositions as well. It is supposed to express the non-distributional character of the hypotheses considered. Indeed, it is the characteristic axiom of what has aptly been called Baconian probability by Cohen (1980), as opposed to Pascalian probability, which is probability proper. (1) and (2) entail that $0 \leq y(A) \leq 1$ for all propositions A. They also entail what I call the *law of negation* (where $\sim A = non\text{-}A$): for all propositions A $\min\{y(A), y(\sim A)\} = 0$. That is, one cannot be potentially surprised by A as well as by $\sim A$. Shackle (1961, ch. IX) argues extensively for that law of negation, but not really for the stronger law of disjunction.

(1) normalizes the maximal degree of potential surprise to be 1. Shackle (1961) leaves open that maximal degree. And since this unit is arbitrary, he claims on p. 135 that potential surprise is measured on a ratio scale. However, his chapter XVI offers only a rudimentary measurement procedure in terms of just noticeable differences in potential or anticipated surprise. He thus follows David Hume’s dubious tendency to assimilate belief or epistemic states in general to more or less strong feelings. In effect, though, the functions of potential surprise remain an ordinal notion.

This also shows in Shackle’s struggle for a corresponding law of conjunction, which takes one form in Shackle (1949, p. 131) repeated in Shackle (1961), but appears in several variants later on. The struggle does not lead to a determinate result. The same holds for his conception of conditional degrees of potential surprise (see Spohn 2012, pp. 228f.).

The Baconian structure was independently reinvented by Rescher (1964) and Cohen (1970) for quite different purposes; Rescher proposed to use it for hypothetical and counterfactual reasoning, whereas Cohen focused on inductive and later, in Cohen (1977), on legal reasoning. Levi (1967) explicitly refers to the work of Shackle, but shifts it to an entirely different place. He conceives of firm beliefs, which leave open so-called serious possibilities. The latter are judged by us in terms of subjective probabilities. (This entails that beliefs, which exclude no serious possibility, have probability 1.) However, the beliefs may change as well; they are, in Levi's peculiar terms, infallible, but corrigible. Levi then uses the functions of potential surprise for accounting for this change. Thus, Levi also assumes a strict division, though not of reality, but in our epistemic constitution.

Another reinvention of the Baconian structure is found in possibility theory, as first displayed in book form in Dubois and Prade (1988). On p. 8 they axiomatically define *possibility measures* Π in such a way that $\Pi(A) = 1 - y(A)$ for some function y of potential surprise. Thus, no potential surprise is maximal possibility, and maximal potential surprise is minimal possibility or impossibility. On pp. 9f. they also introduce the dual notion of degrees of necessity by defining $N(A) = 1 - \Pi(\sim A) = y(\sim A)$. Obviously, one can take any of the three notions as basic and define the other ones.

Dubois and Prade come from an entirely different direction. They are inspired by Lotfi Zadeh's fuzzy logic and in particular by Zadeh (1978) who has already suggested that such measures might be useful. That is, they are inspired by the idea that theories of uncertainty may be subsumed under theories of fuzziness or vagueness. This seems to me to be a serious intuitive confusion. Vagueness is primarily a semantic, not an epistemological phenomenon. For instance, we are uncertain whether Konstanz is a town or a city, because the terms "town" and "city" are semantically vague and Konstanz is a borderline case. This semantic uncertainty is not eliminable; it is not epistemic uncertainty that may be mitigated or even removed by more information. Vagueness is still a hotly discussed topic within philosophy, but along entirely different lines (cf., e.g., Sorensen 2012).

In fact, though, Dubois and Prade are not so much concerned with interpretation, but rather with computability. "Possibility and necessity measure" is just a neutral term for anything that fits this formal structure, just as "necessity" and "possibility" are used in philosophy as umbrella terms for any kind of modality. However, this interpretational indeterminateness also has costs. One is that there is no clear measurement theory for those measures, because it is unclear which magnitude is to be measured by them. Another is that there is no clear guidance as to defining something like conditional degrees of possibility, a point to be emphasized below.

As I shall argue in more detail in Sect. 3, ranking theory can be understood as another reinvention of the Baconian structure. Therefore it is instructive to see how this structure relates to other accounts of modeling epistemic uncertainty. According to Halpern (2003, cf. section 2.8), the most general uncertainty structure consists in so-called *plausibility measures*. These are functions from an algebra of propositions into some set of plausibility values, which need not consist of numbers, but may be any set only partially ordered by $\leq$ and containing $\top$ and $\bot$ as the maximal and the minimal plausibility value. This allows that the ensuing plausibility ordering between propositions contains incomparabilities. Then a plausibility measure Pl has to satisfy the following two axioms:

(3) $\text{Pl}(W) = \top$ and $\text{Pl}(\emptyset) = \bot$,
(4) if $A \subseteq B$, then $\text{Pl}(A) \leq \text{Pl}(B)$ (*monotonicity*).

Of course, it is commonly required that those values are indeed linearly ordered so that we can identify them with real numbers in the interval [0, 1], i.e., $\top = 1$ and $\perp = 0$. Then we arrive at what is called a fuzzy measure by Sugeno (1977) or a Sugeno measure in reference to him and what Schmeidler (1989) calls a non-additive probability.

Now, there are many ways to introduce more special structures. A popular approach to represent uncertainty is as a *set of probability measures* (between which one is undecided); this set may or may not be assumed to be convex. Equivalently, one can represent uncertainty by a so-called *lower probability* which is defined as the pointwise infimum of a set of probabilities and which may also be directly characterized by a somewhat intransparent axiom (cf. Halpern 2003, p. 31). A special case of this is a so-called *inner probability* that is the lower probability of the set of probability measures that extend a given measure on a subalgebra to the full algebra of propositions. Again, inner probabilities can also be directly axiomatized (cf. Halpern, p. 30). In both cases we have the dual notions of *upper* and *outer probabilities.*

Another well-established approach to represent uncertainty consists in *D*(empster)-*S*(hafer) *belief functions*, the dual of which are the *DS plausibility functions* (not to be confused with the plausibility measures above). They seem little known, though, in economic contexts. DS belief functions are axiomatized directly, but such a function can also be understood as a special case of a lower probability, namely the lower probability of the set of probability measures that are pointwise at least as great as that belief function. Halpern (2003, p. 33) emphasizes that there is a (syntactical) sense in which DS belief functions are even equivalent to inner measures. The intended interpretation of DS belief functions is, however, a different one, namely as delivering what is claimed to be an adequate theory of evidence. In particular, the intention is to explicate Keynes' idea of a weight of argument or weight of evidence, which is tentatively put forward in Keynes (1921, ch. VI and pp. 313f.). I am critical of this intention. However, here is not the place for discussing it. Evidence is a deeply problematic notion, which is extensively, but controversially discussed in philosophical epistemology. Undeservedly, the DS theory is little considered there, not even for critical purposes.

Now, necessity and possibility measures are in turn special cases of DS belief and plausibility functions. Shafer (1976, ch. 10) defines and discusses so-called consonant DS belief functions, which are informally characterized as representing consonant evidence, i.e., evidence all "pointing in a single direction" (p. 219). It is easy to show then that possibility measures are nothing but consonant DS plausibility functions. In fact, Shafer (1976, p. 43) also defines so-called degrees of doubt, which relate to his plausibility functions just as Shackle's functions of potential surprise relate to possibility measures, and shows (p. 224) that the functions of potential surprise are just consonant degrees of doubt.

This may suffice as a very sketchy outline of the large field of possible uncertainty representations. In this perspective, functions of potential surprise seem to occupy a small and special place. They do so due to the simple, but quite strong or restrictive Baconian axiom (2). Yet, why precisely should uncertainty obey this axiom? We have not heard much of an argument so far. Rather, we seem well advised to attend to the wider field of uncertainty representations just outlined.

3 Ranking Theory

This verdict seems to apply to ranking functions as well, which, as we will see in a moment, may be conceived as a variant of the functions of potential surprise. However, as I want to explain in this section, it would be inappropriate to shove off ranking theory in this way. The crucial point is this:

All the accounts referred to so far amount to representing something that comes in "degrees". It does not make a difference whether we talk about degrees of belief, of uncertainty, of doubt, of plausibility, or of credibility. Shackle also speaks of degrees of disbelief. All of this comes to the same. The fundamental and obvious fact is that our epistemic attitudes come graded or in degrees, and we have countless ways in natural language to express those degrees. So, in this perspective we have to deliver a theory, or perhaps several theories, of how those degrees behave, no matter how we call them.

Emphatically, though, we also have the notion of *belief*, we talk of it all the time in various ways, and whenever we assert something in an unqualified way—that's the usual way of assertion—we express a belief, equally unqualified or ungraded. This is just as true as that there are degrees of belief.

However, when I look into the literature, at almost every place the notion of belief is used with scare quotes, as if talking of something illegitimate. Traditional epistemologists, searching to improve the old equation "knowledge = justified true belief", have no qualms talking of belief. Formal epistemologists, however, invent all kinds of adjectives, "full, strong, weak, plain, etc. belief", as if this would move things from something forbiddingly unclear to something clearer. And the economic community, just as the AI community, hardly speaks of belief at all. They like to speak of knowledge, but they mean (true) belief and ignore the vigorous philosophical discussion about the nature of knowledge in the last 50 years, which amply displays that knowledge is certainly not just true belief. So, I notice that economists tend to associate certainty or probability 1 with belief, just as with knowledge. The idea seems to be that, when moving to uncertainty (in the embracive sense including risk), we leave the field of belief and enter the field of degrees of belief. However, this idea is definitely misguided. It is very common to take something to be true without being certain of it.

Surely, everybody feels that belief is somehow vague. Consider the dialogue: "Do you believe that Merkel will again be elected chancellor?"—"Yes, I do."—"Do you *really* believe this?"—Well, I am unsure. What do you mean with 'really believe'?" So, aren't we back at degrees of belief in this dialogue? I believe that Merkel will continue as chancellor, but I don't believe it very firmly. Still, the point is that it is correct to say that I believe it.

Ranking theory is precisely made to account for both, belief *and* degrees of belief. This is the virtue that distinguishes it from all the accounts of uncertainty mentioned above. So, what is it? κ is defined to be a *negative ranking function* (for the possibility space W) iff it maps all propositions into the set of natural numbers plus infinity ∞ such that for all disjoint propositions A and B:

(5) $\kappa(W) = 0$ and $\kappa(\emptyset) = \infty$,
(6) $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$ (*the law of disjunction*).

This again entails the above law of negation. Alternatively, we may assume the range of a ranking function to consist in the nonnegative reals (plus infinity) or even in ordinal numbers. There is an argument that one should extend (6) to infinite disjunctions without weakening the minimum to the infimum. This enforces the range to be well-ordered, and then the natural numbers are the more natural choice. Let's neglect the issue here (but see the discussion in Spohn 2012, pp. 72ff.).

So far, (5) and (6) are nothing but a rescaling of functions of potential surprise as defined by (1) and (2), or an inversion and rescaling of possibility measures. What's the point? The point lies in the unequivocal interpretation. Negative ranks (which are non-negative numbers) express degrees of disbelief (that's why they are called negative). Shackle would talk the same way. However, a proposition is disbelieved iff it is taken to be false, and a proposition is believed iff it is taken to be true, i.e., iff its negation is taken to be false. Hence, $\kappa(A) > 0$ means that A is taken to be false, $\kappa(\sim A) > 0$ means that A is believed or taken to be true, and $\kappa(A) = \kappa(\sim A) = 0$ means suspense of judgment with respect to A. In this interpretation, the law of negation states the consistency of beliefs; one can not take both A and $\sim A$ to be false, or to be true. This is a fundamental requirement of rationality.

Clearly, we can also introduce the dual of negative ranking functions. β is a *positive ranking function* iff $\beta(A) = \kappa(\sim A)$ for some negative ranking function κ. Then $\beta(A) > 0$ represents directly that A is believed. It is often convenient to consider *two-sided ranking functions* τ defined by $\tau(A) = \kappa(\sim A) - \kappa(A) = \beta(A) - \kappa(A)$ for some negative ranking function κ. Obviously, we then have $\tau(A) > 0$ or < 0 or $= 0$ according to whether A is taken to be true or taken to be false or neither. Similar notions could have been introduced relative to functions of potential surprise and possibility measures, although the possibilistic analogue of two-sided ranking functions would look very unintuitive.

An important observation is that we may also have stricter notions of disbelief and belief and say that A is disbelieved iff $\kappa(A) > z$ and that A is believed iff $\kappa(\sim A) > z$, for some threshold z; and we could accordingly adapt the interpretation of positive and two-sided ranking functions. This would expand the range of suspense of judgment. And it would precisely account for the vagueness of belief observed above. In this interpretation, (6) displays its full strength by entailing not only the consistency of propositional belief, but also its closure under conjunction and hence its deductive closure, whatever the threshold z. This is another fundamental requirement of rationality. Or it has at least been considered to be so since the establishment of doxastic and epistemic logic by Hintikka (1962).

It is a common objection that requiring deductive closure of beliefs is much too strong an idealization. Let's not enter this delicate issue. However, we should be aware that the idealization lies already in the assumption that our epistemic attitudes relate us to propositions and not to sentences. Hence, all accounts of uncertainty mentioned above suffer from the same idealization. For instance, assigning the same probability (or, e.g., possibility) value to all logically equivalent sentences expressing the same proposition is no easier a task than preserving consistency of beliefs. So, my insistence on those rationality requirements is in no way more problematic than the presuppositions of our entire debate. (One may also argue that these problems arise only with infinite possibility spaces, which can be avoided in many or even most applications.)

In this perspective, the axioms of ranking theory have a clear justification, in contrast to functions of potential surprise and to possibility measures. Shackle sometimes sounds as having the same justification in mind, but often he appears to be appealing to some alleged laws about the feeling of surprise. And regarding something like degrees of possibility it is hard to have intuitions, anyway.

In this way, ranking theory also opens up an epistemological dimension that is simply foreign to all the uncertainty literature. Hence, it is no surprise that the origins of ranking theory lie elsewhere, namely in so-called belief revision theory, which has emerged already in the late 70s and the early state of which is beautifully summarized in Gärdenfors (1988). This theory added a dynamic dimension to the static laws of belief treated by Hintikka (1962), which essentially consist in the consistency and deductive closure of belief. This dynamic dimension studied three kinds of changes of deductively closed belief sets: expansions, contractions (giving up prior beliefs) and revisions (replacing prior beliefs by their negations). The latter two changes were analyzed in terms of a so-called entrenchment ordering. The idea was, roughly, to start deleting the least entrenched beliefs, and then the second least, etc., till the beliefs to be contracted were deleted or, respectively, till the new beliefs could be consistently added. Such an entrenchment ordering sufficed to account for one such change. It was, however, unable to account for iterated belief change. I proposed ranking theory precisely in order to solve this problem. Below I shall indicate how. Essentially, though, the solution was that ranking functions offer a cardinal entrenchment measure and not only an entrenchment ordering. Only later I noticed that the axiom (6) reinvents the Baconian structure already to be found in the work of Shackle and Cohen and, contemporaneously, of Dubois and Prade.

One may object now that simply rescaling the functions and talking in different ways about them does not make a substantial difference. This is true. However, what is so far only a promise is in fact developed to a full-blown theory. Let me indicate only the most important points.

One issue is that degrees of beliefs represented by ranks seem arbitrary. What does it mean to disbelieve something, say, to degree 5 rather than 8? This is answered by a precise measurement theory that delivers a precise operationalization of ranks; see Hild and Spohn (2008). There are various measurement methods. The most interesting one uses as operational base that we can at least state what the beliefs of a person are at any moment (e.g., by asking her). And then it very roughly proceeds thus: You tell me what your beliefs are after any two belief changes, and if your responses are coherent (in a specific sense), I tell you what your determinate ranks are, up to a multiplicative constant. That is, your responses measure your ranks on a ratio scale. The vagueness of belief stated above is then reflected in the fact that the unit of this scale can only be fixed arbitrarily.

This measurement theory is, in a way, analogous to the von Neumann-Morgenstern utility theory. The effects, though, were not analogous. In the utility case the vNM theory immediately decided the debate between ordinalists and cardinalists in favor of the latter. In the belief case a similar effect failed, perhaps because belief revision theory had already developed an impressive corpus and because the measurement theory for ranks was provided only 25 years after their invention. The analogy extends; in the belief case we also have a problem of interpersonal comparison of belief, resembling the vexed issue of the interpersonal comparison of utility. However, this problem has not yet been perceived in the literature. I suspect that it is of similar significance as the problem with utilities.

In any case, this measurement method stays purely on the epistemological side; it does not refer to bets or preferences or similar things. Still, it works and thus adds something not provided elsewhere in the Baconian paradigm since Shackle.

Obviously, the method can only work, if ranking theory also deals with the rational dynamics of belief. This is the second and even more important issue. The standard procedure for stating the dynamics of some mental attitude is to assume a conditional version of that attitude. For instance, the dynamics of subjective probabilities is given by Bayes' theorem and possibly more sophisticated conditionalization rules that refer to conditional probabilities. So, if we want to copy this for ranks, we need to define conditional ranks:

(7) $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$, provided that $\kappa(A) < \infty$, or equivalently:
(8) $\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A)$ (*the law of conjunction*).

I mentioned that Shackle struggled in vain with a suitable law of conjunction and with conditional potential surprise. In Shackle (1961, p. 205) he even considered the analogue of (8), but rejected it, perhaps because of his ordinalist preconceptions. Clearly, using arithmetical operations in (7) and (8) only makes sense if κ represents a cardinal magnitude.

(8) is intuitively convincing. It asks: what's your degree of disbelief in $A \cap B$? One way for $A \cap B$ to be false is that A is false; this contributes $\kappa(A)$ to that degree. However, given A is true, B must be false; so, this adds $\kappa(B \mid A)$ to $\kappa(A \cap B)$. Note also that, given (7), (6) is equivalent to the *conditional law of negation*: $\kappa(B \mid A) = 0$ or $\kappa(\sim B \mid A) = 0$, for all A and B with $\kappa(A) < \infty$. So, the fundamental requirement of ranking theory is just the consistency of conditional belief under all entertainable conditions (i.e., conditions of finite rank). This provides strong normative foundations of ranking theory.

(5)–(8) suggest a translation of probability into ranking theory: replace the sum of probabilities by the minimum of the corresponding ranks and the multiplication and division of probabilities by the addition and subtraction of ranks. This translates the probability axioms into ranking axioms and indeed almost any probabilistic theorem into a corresponding ranking theorem. In each case this ranking theorem makes good sense. I can't explain here the "almost", but see Spohn (2012, pp. 203f.). This indicates that ranking theory indeed unfolds into a rich theory. Note, moreover, that this translation works only in terms of negative ranking functions. This is why they are my preferred theoretical tool.

(7) is the crucial point in which ranking theory transcends the existing Baconian paradigm. Indeed vastly, because it fully opens up the dynamic perspective which had been treated with utter neglect before. It should be relatively clear how to state a general and satisfying dynamics of ranks and thus of beliefs on the basis of conditional ranks. The first idea is to copy probabilistic conditionalization: if A is the content of your experience, then move from κ to κ' with $\kappa'(B) = \kappa(B \mid A)$. This is problematic because it treats experience as maximally certain. However, on this basis more satisfactory conditionalization rules may be stated, which also comprise expansion, revision, and contraction as treated in belief revision theory (cf. Spohn 2012, sect. 5.3–5). This was my original goal.

Let me emphasize the general importance of the topic for accounts of uncertainty. Wherever we study some subject matter, we must study not only its statics, its synchronic laws, but also its dynamics, its diachronic laws. This also holds for epistemology. Hence, it will not do just to state the synchronic form of representations of uncertainty. We must account for their change as well. Probability theory is a shining example, with sophisticated and general rules of change and ample justifications for those rules. This unfolds in detail in the theory of Bayes nets. Ranking theory is another shining example. Indeed, the theory of Bayes nets works just as well in ranking terms (cf. Spohn 2012, ch. 7). What about the other accounts listed in the previous section?

Halpern (2003, chs. 3–4) provides rich answers. He even has an account of conditionalizing his most general plausibility measures. It is clear, though, that all the accounts of uncertainty that can be interpreted in terms of sets of probability measures—most of the accounts mentioned in Sect. 2 are of this kind—are automatically endowed with an account of conditionalization, simply by conditionalizing each measure in the set. Thus, the merits of probabilistic conditionalization carry over to those accounts of uncertainty. However, one would have to check the meaningfulness of this procedure in detail.

Some of the accounts, though, come with an independent idea of conditionalization. I already mentioned that Shackle almost succeeded and in effect foundered. Possibility theory does not provide intuitive access to something like conditional degrees of possibility. Hence, Dubois and Prade (1988) first proposed some definition of conditional possibility that appears quite ad hoc. Later on, they studied also the possibilistic analogue to (7), which works, of course. So, there are various formal options in possibility theory without much interpretational guidance.

The Dempster–Shafer theory also comes with a full-fledged dynamic account, based on Dempster's rule of combination. This was always perceived as one of its strengths. I mentioned in Sect. 2 that ranking theory seems to reduce to the DS theory, since ranking functions, like possibility measures, appear to be nothing but consonant DS belief or plausibility functions. In fact, though, things do not fit. A DS belief function learns uncertain evidence by getting combined with a so-called simple support function. A ranking function does so according to some form of conditionalization. The point now is that the combination of a consonant belief function with a simple support function need no longer be consonant. Thus, the dynamic rules diverge, and the theory reduction fails. I have given this argument in Spohn (1990). It shows more generally that the idea of consonant belief functions is entirely foreign to ranking theory; there need not be anything ‘consonant’ in a ranking function. If one conceives of DS belief functions as lower probabilities, i.e., as lower bounds of sets of probability measures, the above probabilistic strategy leads to other definitions of conditional DS belief functions (cf. Halpern 2003, pp. 92ff.). However, this does not change the situation; ranking theory does still not reduce to the DS theory; see Spohn (2012, pp. 266ff.).

The upshot of these brief remarks is that those dynamic considerations confirm the meaningfulness, the strength, and the independence of ranking theory. This was the point I wanted to display.

One may doubt this independence on different grounds. Above I sketched a translation of probability into ranking theory. This translation suggests that negative ranks are just something like the logarithm of probabilities relative to a small base. Therefore, negative ranks are sometimes called order-of-magnitude probabilities. This formal analogy may be helpful. However, it is also very misleading. The logarithmic transformation works perfectly only if the base is infinitesimal. Then disbelief would be very small or even infinitesimal probability, and belief would be very large probability or probability $1 - i$ for some infinitesimal i. However, phenomenologically this is just wrong; beliefs need not be (almost) certain. I would bet my life on an event of probability $1 - i$, but not on most of my beliefs. And it founders at the lottery paradox which shows that high probability does not behave like (deductively closed) belief (see, e.g., Wheeler 2007).

What then is the relation between probabilities and ranks? I don’t know. One cannot separate them by confining them to different kinds of propositions, as Knight and Shackle and even Levi seem to have suggested. They both apply to all propositions. It does not seem possible to reduce beliefs or even ranks to probabilities, despite vigorous efforts. I am referring here to the rich literature on the lottery paradox (again see Wheeler 2007), which leaves me unsatisfied. Leitgeb (2017) attempts to reconcile probability and belief, at the cost of what appears to me to be an intolerable partition-dependence of belief. There are formal unifications, but it is unclear whether they have more than formal meaning. I discuss this in Spohn (2012, ch. 10), and my preliminary conclusion is what I call methodological separatism, which recommends accepting and developing both theories independently, as long as there is no convincing account of their relation. The strong motive behind this recommendation is, however, that ranking theory provides such a rich theory of belief that it cannot be dismissed or marginalized. We are challenged to take a stance on that relation.

4 Decision Theory Based on Uncertainty?

So far, all of this was pure epistemology: how to adequately describe epistemic states and their dynamics? I have sketched this a bit more broadly in order to emphasize that there is substantial theorizing that is fruitful by itself and not dependent on an embedding in a practical or behavioral theory or a theory of rational decision or action. If that embedding does not run smoothly, it is an open question where the problem lies: is the epistemology to be blamed or are the attempts to construct a suitable decision theory insufficient? After these precautionary remarks I admit, of course, that economists are predominantly interested precisely in this embedding. This was Knight's motive, as it was Shackle's. So, let's see whether the uncertainty measures discussed so far are accompanied by a decision-theoretic extension.

Without doubt, the unmatched paradigm is standard decision theory, which combines probability and a vNM utility function in order to state the Bayesian principle of maximizing expected utility. This is developed in an exemplary way by Savage (1954). There is surprisingly subtle variation within that paradigm. E.g., since more than 40 years philosophers make much ado about the opposition between what they call causal and evidential decision theory, while this is hardly reflected in economics (see, e.g., Weirich 2016). Still, the paradigm is relatively unified. What's the offer of the other uncertainty measures?

Well, there is the probabilistic strategy. We observed above that many uncertainty measures are associated with or determined by sets of probability measures. And thus each act receives an entire set of standard expected utilities, also representable by a lower and an upper expected utility. This raises the issue which decision rule to accept, which of the many expected utilities to maximize. And then the old discussions about decisions under uncertainty get revived in more sophisticated form. Again, we can consider maximin and maximax rules and further variants. (See, e.g., Halpern 2003, ch. 5.)

Apart from this, there was silence, initially. Shackle did not arrive at a determinate proposal. And the literature on Baconian probability did not connect up at all. As far as ranking theory is concerned, I sensed the gap, but I was content with the epistemological applications of ranking theory in philosophy of science, in particular in the theory of deterministic causation.

However, people became soon aware that one could use the theory of capacities of Choquet (1953) in order to develop a theory of integration or expectation for other uncertainty measures directly, i.e., not via the probabilistic strategy. Dempster (1967) had such an idea for DS belief functions, and Sugeno (1977) did so for fuzzy measures (although his fuzzy integral differs from the Choquet integral). As far as I see, though, these explorations did not yet have a specific decision theoretic content.

The situation changed with Schmeidler (1989) and Gilboa (1987), who ingeniously copied the procedure of Savage (1954). Start with a preference relation over a rich set of acts, assume that it satisfies certain axioms (weaker than the Savage axioms) and prove that there is a unique non-additive probability measure (in Schmeidler's sense) and a vNM utility function measured on an interval scale such that the preference relation between acts matches the expected utility of acts—where expectation with respect to a non-additive probability is construed in terms of the Choquet integral. So, the Savage kind of justification of subjective probabilities and utilities is amenable also to other epistemic formats. Hence, it is no surprise that variants of this strategy have been applied to other uncertainty measures. Dubois et al. (2001), e.g., prove similar theorems for possibility measures.

Giang and Shenoy (2000) took a different route for ranking functions, which, as far as I see, cannot be subsumed under Schmeidler's strategy. They proceeded from the von Neumann-Morgenstern axioms for preferences about lotteries, replaced probabilistic by ranking lotteries, and arrived thus at analogous results. Let me briefly indicate how this works:

We start with an exhaustive set $C = \{c_1, \ldots, c_n\}$ of mutually exclusive consequences or outcomes. "Lottery" sounds very probabilistic. Let's better talk of prospects (even though this term is occupied by still another theory). A *prospect* $p = [r_1 c_1, \ldots, r_n c_n]$ then is a mixture of the outcomes c_i by the negative ranks $\kappa(c_i) = r_i$ ($i = 1, \ldots, n$). So, at least one of the r_i must be 0. Note that such a prospect can be informatively specified only in terms of negative ranks. This is so because the outcomes are mutually exclusive so that at most one of them can be positively believed to obtain, while all others receive positive rank 0, even though they can differ in negative ranks.

The idea now is to assign to each prospect p a disutility/utility pair $DU(p) = \langle x, y \rangle$, where x and y are natural numbers or ∞. The first member x is the disutility of p and the second member y the utility of p. Hence, either $x = 0$ or $y = 0$, since a prospect cannot have both, a positive disutility and a positive utility. How does this assignment work?

The first step is to assume (without loss of generality) that c_1 is the (or a) best and c_n the (or a) worst outcome. Then we can define a *standard prospect* $[r_1 c_1, r_n c_n]$ as one only between the best and the worst outcome, where either $r_1 = 0$ or $r_n = 0$ or both (the other outcomes are excluded by $r_i = \infty$ for $i \neq 1, n$). These standard prospects correspond to the well-known BRLT's (basic reference lottery tickets) of Raiffa (1968).

Such a standard prospect $p = [r_1 c_1, r_n c_n]$ is to receive the disutility/utility pair $DU(p) = \langle r_1, r_n \rangle$: If $r_1 = 0$, the utility of p is r_n = the negative rank of the worst outcome c_n = the positive rank of the best outcome c_1; the disutility of p is r_1 = the negative rank of c_1 = the positive rank of c_n. That's quite intuitive, and indeed the only assignment for which the construction works. Note that thereby such disutility/utility pairs are measured on an absolute scale with a fixed zero and a fixed unit. Recall, however, that ranks were measured only on a ratio scale with a fixed zero and an arbitrary unit. Hence the utilities covary with the ranks. This differs from what we are used from probabilities and vNM utilities.

The second step is to assume that each outcome c_i is equivalent to some standard prospect $p = [r_1 c_1, r_n c_n]$. Thereby each outcome c_i also receives the disutility/utility pair $DU(c_i) = \langle r_1, r_n \rangle$. This entails that $DU(c_1) = \langle 0, \infty \rangle$ and $DU(c_n) = \langle \infty, 0 \rangle$, i.e., that c_1 has infinite utility and c_n infinite disutility. This consequence seems embarrassing. However, it does not express that c_1 (and c_n) cannot be outweighed; it only says that in a standard prospect $[r_1 c_1, r_n c_n]$ you get c_1 for sure only if $r_n = \infty$.

The third step is to extend the construction to all prospects. Our plan is the same as in vNM utility theory. Replace each outcome in a prospect by its standard prospect and determine then how firmly you (dis)believe to get the best or, respectively, the worst outcome in the resulting compound prospect. More precisely, for any prospect $p = [r_1 c_1, \ldots, r_n c_n]$ and any outcome c_i with $DU(c_i) = \langle x, y \rangle$ define the weighted disutility/utility pair of c_i weighted by r_i as $WDU(c_i, r_i) = \langle x + r_i, y + r_i \rangle$. Observe that $WDU(c_i, r_i)$ need not be a disutility/utility pair, since both its members may be positive. Finally, define the *disutility/utility pair* of p as $DU(p) = \min_{i \leq n} WDU(c_i, r_i)$, which is the pair of the minimum of all first members and the minimum of all second members of the pairs $WDU(c_i, r_i)$. Then, indeed, at least one of the two minima must be 0. This definition of $DU(p) = \langle x, y\rangle$ realizes our plan: if the c_i in p are replaced by their equivalent standard prospect, then a compound prospect results, in which x and y are, respectively, the negative and the positive rank with which c_1 is expected in that compound prospect. This may be easily derived from the basic laws (6) and (8), the laws of disjunction and conjunction.

The procedure looks a bit complicated. However, first summing (as in $WDU(c_i, r_i)$) and then taking the minimum (as in $DU(p)$) is the pertinent ranking-theoretic analogue to the usual way of averaging vNM utilities by first multiplying them with probabilistic weights and then summing the weighted utilities. Thus, $DU(p)$ is, in a way, an expected disutility/utility pair.

A toy example may illustrate the matter. Suppose you have a moderately satisfying tenured academic job (J). Now you get an offer for a really exciting research position, which, however, is not yet tenured. If you do well, you get tenure after five years (T); if not, you have to look for a new job (N). Clearly, you prefer T to J, but also J to N. What should you do?

Standard prospects are between T and N. Thus, J is equivalent to some such standard prospect. Suppose that $DU(J) = \langle 0, 0\rangle$. This amounts to suspense of judgment between T and N. Hence, if you are confident to succeed in the new job, i.e., $\tau(T) > 0$ (in terms of your two-sided ranks), you should accept the offer; if you are doubtful, i.e., $\tau(T) < 0$, you should decline it; and if you suspend judgment, you may do either way. So far, this is the same as within standard decision theory: if you believe in success firmly enough, you should try.

Now, let's slightly enrich the example. Suppose that in response to the offer you received your present employer grants you access to a promotion program. If you are evaluated positively, you get the promotion (P); if not, you stay in your present job (J). Of course, P is better than J, but not as attractive as T. So, you have the choice between the T/N-prospect and P/J-prospect. Again, what to do?

Clearly, the answer depends on your confidence $\tau(T)$ to succeed in the new job, on your confidence $\tau(P)$ to get the promotion, and on $DU(P) = \langle 0, x\rangle$, the utility of getting the promotion, which is positive, i.e., $x > 0$. Thus, P is equivalent to the standard prospect $[0T, xN]$. Then the above rule for calculating the disutility/utility pair for compound prospects entails that you should accept the new job iff $\tau(T) > \min\ \{x, \tau(P)\}$, i.e., iff either your confidence in getting the promotion is lower than that in succeeding in the new job or the utility of the promotion is lower than that of the uncertain prospect of the new job. Conversely, you should decline the offer iff your confidence in the promotion is higher than that in success in the new job (as it should since the promotion should be easier) *and* the utility of the promotion is higher than that of the uncertain prospect of the new job.

This recommendation diverges from how standard decision theory would represent the case. Still, it results from a kind of weighing beliefs and desires, though a more coarse-grained weighing than the probabilistic one. The recommendation sounds quite plausible in the case presented. I grant that in other examples the recommendations may appear quite implausible. It may be hard to say, though, whether this is due to the fact that our intuitions are strongly shaped by standard decision theory. So much about ranking-theoretic decision theory.

To resume, there is not only the probabilistic strategy of extending decision theory to other epistemic formats through their representation by sets of probability measures. There is also the direct strategy, as proposed by Schmeidler and as just exemplified for ranking theory in a different way. The justificatory business is not really facilitated thereby. Originally, the Savage preference axioms appeared most convincing, despite the criticisms they received. But now we end up with a multitude of different preference axioms, all of which induce varying epistemic formats, utility functions, and decision rules. Thereby, the justificatory burden shifts to those axioms. I find neither much of a normative discussion about those axioms nor attempts of embedding them in a wider reflective equilibrium of normative theorizing. People rather seem to take only an intrinsic interest in stating lots of formal possibilities.

I have another discontent with that literature. I do not know well enough what these alternative decision theories amount to. In the abstract it is clear how to develop them. In principle, all decision-theoretic theorizing can be reduced to the consideration of basic lotteries. Thus, by exchanging the kind of basic lotteries, one could develop the alternatives in parallel. However, to my knowledge this has not been done. In Dubois et al. (2001) no example at all is discussed. Schmeidler (1989) is happy to be able to account for the obstinate preferences in the Ellsberg paradox; this is indeed impressive. Giang and Shenoy (2000) discuss a small example that is a bit more complex than mine above. I am not aware of any more extensive theorizing.

This disappoints me. I was always impressed by the power and by the wide ramifications of standard decision theory, which allowed applying it to complex situations in a meaningful way. How are those situations treated by alternative decision theories? Take any more advanced textbook on statistical decision theory, on planning theory, or what have you, and try to translate it into alternative epistemic formats. Do we generate sensible results thereby? Or consider game theory, an extension that is crucial for economists. How to do game theory in terms of those other uncertainty measures? This is very unclear. What, for instance, could be a Nash equilibrium in mixed strategies in those terms? The reduction of game theory to decision theory is most specifically explored in epistemic game theory. However, I don't see any way of doing epistemic game theory in terms of other uncertainty measures. And so on.

In short, in my view justification also lies in the simple fact of being a rich and substantial theory. Of course, not every idea that is developed into a rich theory is thereby justified. But the space of potential justification is tremendously expanded thereby. This is a value in itself. And I miss this value for all those alternative decision theories. Therefore I well understand economists who do not feel attracted by this field.

However, the adequate response is not to dismiss the field. Recall my precautionary remarks at the beginning of this section. There is ample epistemological reason to consider uncertainty measures besides subjective probabilities. And if we think we need decision-theoretic extensions of those measures, we have to work hard at those extensions instead of discarding those measures when those extensions seem unwieldy.

Let me remark, finally, that in this respect the standing of ranking theory is no better, but also no worse than that of the other uncertainty measures. The appeal to further development applies to ranking-theoretic decision theory no less than to the other measures. However, I have emphasized the epistemological value of ranking theory. This may be a strong motive for particularly engaging in its decision theoretic extension and perhaps for further pursuing the above suggestion of a ranking-theoretic decision theory.

References

Carnap, R. (1971). A basic system of inductive logic, part I. In R. Carnap & R. C. Jeffrey (Eds.), *Studies in inductive logic and probability* (Vol. I, pp. 33–165). Berkeley: University of California Press.

Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Fourier, 5,* 131–295.

Cohen, L. J. (1970). *The implications of induction*. London: Methuen.

Cohen, L. J. (1977). *The probable and the provable*. Oxford: Oxford University Press.

Cohen, L. J. (1980). Some historical remarks on the Baconian conception of probability. *Journal of the History of Ideas, 41,* 219–231.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics, 38,* 325–339.

Dubois, D., & Prade, H. (1988). *Possibility theory: An approach to computerized processing of uncertainty*. New York: Plenum Press.

Dubois, D., Prade, H., & Sabbadin, Régis. (2001). Decision-theoretic foundations of qualitative possibility theory. *European Journal of Operational Research, 128,* 459–478.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economies, 75,* 645–669.

Gärdenfors, P. (1988). *Knowledge in flux. Modeling the dynamics of epistemic states*. Cambridge: MIT Press.

Giang, P. H., & Shenoy, P. P. (2000). A qualitative linear utility theory for Spohn's theory of epistemic beliefs. In C. Boutilier & M. Goldszmidt (Eds.), *Uncertainity in artificial intelligence* (Vol. 16, pp. 220–229). San Francisco: Morgan Kaufmann.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics, 16,* 65–88.

Halpern, J. Y. (2003). *Reasoning about uncertainty*. Cambridge: MIT Press.

Hild, M., & Spohn, W. (2008). The measurement of ranks and the laws of iterated contraction. *Artificial Intelligence, 172,* 1195–1218.

Hintikka, J. (1962). *Knowledge and belief.* Ithaca: Cornell University Press.

Keynes, J. M. (1921). *A treatise on probability*. London: Macmillan.

Knight, F. H. (1921). *Risk, uncertainty, and profit*. Boston: Hart, Schaffner & Marx.

Leitgeb, H. (2017). *The stability of belief: How rational belief coheres with probability*. Oxford: Oxford University Press.

Levi, I. (1967). *Gambling with truth. An essay on induction and the aims of science*. New York: Knopf.

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in inductive logic and probability* (Vol. II, pp. 263–293). Berkeley: University of California Press.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo: Morgan Kaufmann.

Raiffa, H. (1968). *Introduction into decision analysis*. Reading: Addison-Wesley.

Rescher, N. (1964). *Hypothetical reasoning*. Amsterdam: North-Holland.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley. **(2nd ed. in Dover 1972)**.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica, 57,* 571–587.

Shackle, G. L. S. (1949). *Expectation in economics*. Cambridge: Cambridge University Press.

Shackle, G. L. S. (1961). *Decision, order, and time in human affairs*. Cambridge: Cambridge University Press. **(2nd ed. in 1969)**.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton: Princeton University Press.

Shafer, G. (1978). Non-additive probabilities in the work of Bernoulli and Lambert. *Archive for History of Exact Sciences, 19,* 309–370.

Sorensen, R. (2012). Vagueness. In *Stanford encyclopedia of philosophy*. https://plato.stanford.edu/entries/vagueness/.

Spohn, W. (1988). Ordinal conditional functions. A dynamic theory of epistemic states. In W. L. Harper & B. Skyrms (Eds.), *Causation in decision, belief change, and statistics* (Vol. II, pp. 105–134). Dordrecht: Kluwer.

Spohn, W. (1990). A general non-probabilistic theory of inductive reasoning. In R. D. Shachter, T. S. Levitt, J. Lemmer, & L. N. Kanal (Eds.), *Uncertainty in artificial intelligence* (Vol. 4, pp. 149–158). Amsterdam: Elsevier.

Spohn, W. (2012). *The laws of belief. Ranking theory and its philosophical applications*. Oxford: Oxford University Press.

Sugeno, M. (1977). Fuzzy measures and fuzzy integrals—A survey. In M. M. Gupta et al. (Eds.), *Fuzzy automata and decision processes* (pp. 89–102). Amsterdam: North-Holland.

Weirich, P. (2016). Causal decision theory. In *Stanford encyclopedia of philosophy*. https://plato.stanford.edu/entries/decision-causal/.

Wheeler, G. (2007). A review of the lottery paradox. In W. L. Harper & G. Wheeler (Eds.), *Probability and inference: Essays in honour of Henry E. Kyburg, Jr.* (pp. 1–31). London: King's College Publications.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems, 1,* 3–28.