# Chapter 17
# A Survey of Ranking Theory

**Wolfgang Spohn**

## Introduction

Epistemology is concerned with the fundamental laws of thought, belief, or judgment. It may inquire the fundamental relations among the objects or contents of thought and belief, i.e., among propositions or sentences. Then we enter the vast realm of formal logic. Or it may inquire the activity of judging or the attitude of believing itself. Often, we talk as if this would be a yes or no affair. From time immemorial, though, we know that judgment is firm or less than firm, that belief is a matter of degree. This insight opens another vast realm of formal epistemology.

Logic received firm foundations already in ancient philosophy. It took much longer, though, until the ideas concerning the forms of (degrees of) belief acquired more definite shape. Despite remarkable predecessors in Indian, Greek, Arabic, and medieval philosophy, the issue seemed to seriously enter the agenda of intellectual history only in the sixteenth century with the beginning of modern philosophy. Cohen (1980) introduced the handy, though somewhat tendentious opposition between Baconian and Pascalian probability. This suggests that the opposition was already perceivable with the work of Francis Bacon (1561–1626) and Blaise Pascal (1623–1662). In fact, philosophers were struggling to find the right mould. In that struggle, Pascalian probability, which *is* probability *simpliciter*, was the first to take a clear and definite shape, viz. in the middle of seventeenth century (cf. Hacking 1975), and since then it advanced triumphantly. The extent to which it interweaves with our cognitive enterprise has become nearly total (cf. the marvelous collection of Krüger et al. 1987). There certainly were alternative ideas. However, probability theory was always far ahead; indeed, the distance ever increased. The winner takes it all!

W. Spohn (✉)
Fachbereich Philosophie, Universität Konstanz, 78457 Konstanz, Germany
e-mail: wolfgang.spohn@uni-konstanz.de



H. Arló-Costa et al. (eds.), *Readings in Formal Epistemology*, Springer Graduate Texts in Philosophy 1, DOI 10.1007/978-3-319-20451-2_17

I use ‘Baconian probability’ as a collective term for the alternative ideas. This is legitimate since there are strong family resemblances among the alternatives. Cohen has chosen an apt term since it gives historical depth to ideas that can be traced back at least to Bacon (1620) and his powerful description of ‘the method of lawful induction’. Jacob Bernoulli and Johann Heinrich Lambert struggled with a non-additive kind of probability. When Joseph Butler and David Hume spoke of probability, they often seemed to have something else or more general in mind than our precise explication. In contrast to the German Fries school British nineteenth century’s philosophers like John Herschel, William Whewell, and John Stuart Mill elaborated non-probabilistic methods of inductive inference. And so forth.[1]

Still, one might call this an underground movement. The case of alternative forms of belief became a distinct hearing only in the second half of the twentieth century. On the one hand, there were scattered attempts like the ‘functions of potential surprise’ of Shackle (1949), heavily used and propagated in the epistemology of Isaac Levi since his (1967), Rescher’s (1964) account of hypothetical reasoning, further developed in his (1976) into an account of plausible reasoning, or Cohen’s (1970) account of induction which he developed in his (1977) under the label ‘Non-Pascalian probability’, later on called ‘Baconian’. On the other hand, one should think that modern philosophy of science with its deep interest in theory confirmation and theory change produced alternatives as well. Indeed, Popper’s hypothetical-deductive method proceeded non-probabilistically, and Hempel (1945) started a vigorous search for a qualitative confirmation theory. However, the former became popular rather among scientists than among philosophers, and the latter petered out after 25 years, at least temporarily.

I perceive all this rather as a prelude, preparing the grounds. The outburst came only in the mid 70s, with strong help from philosophers, but heavily driven by the needs of Artificial Intelligence. Not only deductive, but also inductive reasoning had to be implemented in the computer, probabilities appeared intractable[2], and thus a host of alternative models were invented: a plurality of default logics, non-monotonic logics and defeasible reasonings, fuzzy logic as developed by Zadeh (1975, 1978), possibility theory as initiated by Zadeh (1978) and developed by Dubois and Prade (1988), the Dempster-Shafer belief functions originating from Dempster (1967, 1968), but essentially generalized by Shafer (1976), AGM belief revision theory (cf. Gärdenfors 1988), a philosophical contribution with great success in the AI market, Pollock’s theory of defeasible reasoning (summarized in Pollock 1995), and so forth. The field has become rich and complex. There are attempts of unification like Halpern (2003) and huge handbooks like Gabbay et al. (1994). One hardly sees the wood for trees. It seems that what had been forgotten for centuries had to be made good for within decades.

---

[1]This is not the place for a historical account. See, e.g., Cohen (1980) and Shafer (1978) for some details.

[2]Only Pearl (1988) showed how to systematically deal with probabilities without exponential computational explosion.

Ranking theory, first presented in Spohn (1983, 1988)[3], belongs to this field as well. Since its development, by me and others, is scattered in a number of papers, one goal of the present paper is to present an accessible survey of the present state of ranking theory.[4] This survey will emphasize the philosophical applications, thus reflecting my bias towards philosophy. My other goal is justificatory. Of course, I am not so blinded to claim that ranking theory would be *the* adequate account of Baconian probability. As I said, 'Baconian probability' stands for a collection of ideas united by family resemblances; and I shall note some of the central resemblances in the course of the paper. However, there is a multitude of epistemological purposes to serve, and it is entirely implausible that there is one account to serve all. Hence, postulating a reign of probability is silly, and postulating a duumvirate of probability and something else is so, too. Still, I am not disposed to see ranking theory as just one offer among many. On many scores, ranking theory seems to me to be superior to rival accounts, the central score being the notion of *conditional* ranks. I shall explain what these scores are, thus trying to establish ranking theory as one particularly useful account of the laws of thought.

The plan of the paper is simple. In the five subsections of section "The theory", pp. 305ff, I shall outline the main aspects of ranking theory. This central section will take some time. I expect the reader to get impatient meanwhile; you will get the strong impression that I am not presenting an alternative to (Pascalian) probability, as the label 'Baconian' suggests, but simply probability itself in a different disguise. This is indeed one way to view ranking theory, and a way, I think, to understand its virtues. However, the complex relation between probability and ranking theory, though suggested at many earlier points, will be systematically discussed only in section "Ranks and probabilities", pp. 328ff. The section "Further comparisons", pp. 335ff, will finally compare ranking theory to some other accounts of Baconian probability.

## The Theory

### *Basics*

We have to start with fixing the objects of the cognitive attitudes we are going to describe. This is a philosophically highly contested issue, but here we shall stay conventional without discussion. These objects are pure contents, i.e., propositions.

[3]There I called its objects ordinal conditional functions. Goldszmidt and Pearl (1996) started calling them ranking functions, a usage I happily adapted.

[4]In the meantime, my comprehensive book on ranking theory (Spohn 2012) has appeared. This paper may also serve as an introduction to this book. Reversely, various topics, which are only touched here and then referred back to older papers of mine, are developed in this book in a better and more comprehensive way.

To be a bit more explicit: We assume a non-empty set $W$ of mutually exclusive and jointly exhaustive possible worlds or *possibilities*, as I prefer to say, for avoiding the grand associations of the term 'world' and for allowing to deal with *de se* attitudes and related phenomena (where doxastic alternatives are considered to be centered worlds rather than worlds). And we assume an algebra $\mathcal{A}$ of subsets of $W$, which we call *propositions*. All the functions we shall consider for representing doxastic attitudes will be functions defined on that algebra $\mathcal{A}$.

Thereby, we have made the philosophically consequential decision of treating doxastic attitudes as intensional. That is, when we consider sentences such as "$a$ believes (with degree $r$) that $p$", then the clause $p$ is substitutable salva veritate by any clause $q$ expressing the same proposition and in particular by any logically equivalent clause $q$. This is so because by taking propositions as objects of belief we have decided that the truth value of such a belief sentence depends only on the proposition expressed by $p$ and not on the particular way of expressing that proposition. The worries provoked by this decision are not our issue.

The basic notion of ranking theory is very simple:

**Definition 17.1** Let $\mathcal{A}$ be an algebra over $W$. Then $\kappa$ is a *negative ranking function*[5] for $\mathcal{A}$ iff $\kappa$ is a function from $\mathcal{A}$ into $\mathbf{R}^* = \mathbf{R}^+ \cup \{\infty\}$ (i.e., into the set of non-negative reals plus infinity) such that for all $A, B \in \mathcal{A}$:

$$\kappa(W) = 0 \text{ and } \kappa(\varnothing) = \infty, \tag{17.1}$$

$$\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\} \qquad [\textit{the law of disjunction (for negative ranks)}]\,. \tag{17.2}$$

$\kappa(A)$ is called the (*negative*) *rank* of $A$.

It immediately follows for each $A \in \mathcal{A}$:

$$\text{either } \kappa(A) = 0 \text{ or } \kappa(\overline{A}) = 0 \text{ or both} \qquad [\textit{the law of negation}]\,. \tag{17.3}$$

A negative ranking function $\kappa$, this is the standard interpretation, expresses a *grading of disbelief* (and thus something negative, hence the qualification). If $\kappa(A) = 0$, $A$ is not disbelieved at all; if $\kappa(A) > 0$, $A$ is disbelieved to some positive degree. Belief in $A$ is the same as disbelief in $\overline{A}$; hence, $A$ is *believed* in $\kappa$ iff $\kappa(\overline{A}) > 0$. This entails (via the law of negation), but is not equivalent to $\kappa(A) = 0$. The latter is compatible also with $\kappa(\overline{A}) = 0$, in which case $\kappa$ is neutral or unopinionated concerning $A$. We shall soon see the advantage of explaining belief in this indirect way via disbelief.

A little example may be instructive. Let us look at Tweetie of which default logic is very fond. Tweetie has, or fails to have, each of the three properties: being a bird

[5]For systematic reasons I am slightly rearranging my terminology from earlier papers. I would be happy if the present terminology became the official one.

($B$), being a penguin ($P$), and being able to fly ($F$). This makes for eight possibilities. Suppose you have no idea what Tweetie is, for all you know it might even be a car. Then your ranking function may be the following one, for instance:[6]

| κ | $B\&\overline{P}$ | $B\ \&\ P$ | $\overline{B}\&\overline{P}$ | $\overline{B}\&P$ |
|---|---|---|---|---|
| $F$ | 0 | 4 | 0 | 11 |
| $\overline{F}$ | 2 | 1 | 0 | 8 |

In this case, the strongest proposition you believe is that Tweetie is *either* no penguin and no bird $\left(\overline{B}\&\overline{P}\right)$ *or* a flying bird and no penguin ($F\ \&\ B\ \&\ \overline{P}$). Hence, you neither believe that Tweetie is a bird ($B$) nor that it is not a bird ($\overline{B}$). You are also neutral concerning its ability to fly. But you believe, for instance: if Tweetie is a bird, it is not a penguin and can fly $\left(B \rightarrow \overline{P}\&F\right)$; and if Tweetie is not a bird, it is not a penguin $\left(\overline{B} \rightarrow \overline{P}\right)$ – each if-then taken as material implication. In this sense you also believe: if Tweetie is a penguin, it can fly ($P \rightarrow F$); and if Tweetie is a penguin, it cannot fly $\left(P \rightarrow \overline{F}\right)$ – but only because you believe that it is not a penguin in the first place; you simply do not reckon with its being a penguin. If we understand the if-then differently, as we shall do later on, the picture changes. The larger ranks in the last column indicate that you strongly disbelieve that penguins are not birds. And so we may discover even more features of this example.

What I have explained so far makes clear that we have already reached the first fundamental aim ranking functions are designed for: the *representation of belief*. Indeed, we may define $\mathcal{B}_\kappa = \{A \mid \kappa(\overline{A}) > 0\}$ to be the *belief set* associated with the ranking function κ. This belief set is finitely *consistent* in the sense that whenever $A_1, \ldots, A_n \in \mathcal{B}_\kappa$, then $A_1 \cap \ldots \cap A_n \neq \varnothing$; this is an immediate consequence of the law of negation. And it is finitely *deductively closed* in the sense that whenever $A_1, \ldots, A_n \in \mathcal{B}_\kappa$ and $A_1 \cap \ldots \cap A_n \subseteq B \in \mathcal{A}$, then $B \in \mathcal{B}_\kappa$; this is an immediate consequence of the law of disjunction. Thus, belief sets just have the properties they are normally assumed to have. (The finiteness qualification is a little cause for worry that will be addressed soon.)

There is a big argument about the rationality postulates of consistency and deductive closure; we should not enter it here. Let me only say that I am disappointed by all the attempts I have seen to weaken these postulates. And let me point out that the issue was essentially decided at the outset when we assumed belief to operate on propositions or truth-conditions or sets of possibilities. With these assumptions we ignore the relation between propositions and their sentential expressions or modes of presentation; and it is this relation where all the problems hide.

[6]I am choosing the ranks in an arbitrary, though intuitively plausible way (just as I would have to arbitrarily choose plausible subjective probabilities, if the example were a probabilistic one). The question how ranks may be measured will be taken up in section "The dynamics of belief and the measurement of belief", pp. 316ff.

When saying that ranking functions represent belief I do not want to further qualify this. One finds various notions in the literature, full beliefs, strong beliefs, weak beliefs, one finds a distinction of acceptance and belief, etc. In my view, these notions and distinctions do not respond to any settled intuitions; they are rather induced by various theoretical accounts. Intuitively, there is only one perhaps not very clear, but certainly not clearly divisible phenomenon which I exchangeably call believing, accepting, taking to be true, etc.

However, if the representation of belief were our only aim, belief sets or their logical counterparts as developed in doxastic logic (see already Hintikka 1962) would have been good enough. What then is the purpose of the ranks or degrees? Just to give another account of the intuitively felt fact that belief is graded? But what guides such accounts? Why should degrees of belief behave like ranks as defined? Intuitions by themselves are not clear enough to provide this guidance. Worse still, intuitions are usually tainted by theory; they do not constitute a neutral arbiter. Indeed, problems already start with the intuitive conflict between representing belief and representing degrees of belief. By talking of belief *simpliciter*, as I have just insisted, I seem to talk of *ungraded* belief.

The only principled guidance we can get is a theoretical one. The degrees must serve a clear theoretical purpose and this purpose must be shown to entail their behavior. For me, the theoretical purpose of ranks is unambiguous; this is why I invented them. It is the *representation of the dynamics of belief*; that is the second fundamental aim we pursue. How this aim is reached and why it can be reached in no other way will unfold in the course of this section. This point is essential; as we shall see, it distinguishes ranking theory from all similarly looking accounts, and it grounds its superiority.

For the moment, though, let us look at a number of variants of Definition 17.1. Above I mentioned the finiteness restriction of consistency and deductive closure. I have always rejected this restriction. An inconsistency is irrational and to be avoided, be it finitely or infinitely generated. Or, equivalently, if I take to be true a number of propositions, I take their conjunction to be true as well, even if the number is infinite. If we accept this, we arrive at a somewhat stronger notion:

**Definition 17.2** Let $\mathcal{A}$ be a complete algebra over $W$ (closed also under infinite Boolean operations). Then $\kappa$ is a *complete negative ranking function* for $\mathcal{A}$ iff $\kappa$ is a function from $W$ into $\mathbf{N}^+ = \mathbf{N} \cup \{\infty\}$ (i.e., into the set of non-negative integers plus infinity) such that $\kappa^{-1}(0) \neq \varnothing$ and and $\kappa^{-1}(n) \in \mathcal{A}$ for each $n \in \mathbf{N}^+$. $\kappa$ is extended to propositions by defining $\kappa(\varnothing) = \infty$ and $\kappa(A) = \min\{\kappa(w) \mid w \in A\}$ for each non-empty $A \in \mathcal{A}$.

Obviously, the propositional function satisfies the laws of negation and disjunction. Moreover, we have for any $\mathcal{B} \subseteq \mathcal{A}$:

$$\kappa(\cup\mathcal{B}) = \min\{\kappa(B) \mid B \in \mathcal{B}\} \qquad [\textit{the law of infinite disjunction}]. \qquad (17.4)$$

Due to completeness, we could start in Definition 17.2 with the point function and then define the set function as specified. Equivalently, we could have defined the set functions by the conditions (17.1) and (17.4) and then reduce the set function to a point function. Henceforth I shall not distinguish between the point and the set function. Note, though, that without completeness the existence of an underlying point function is not guaranteed; the relation between point and set function in this case is completely cleared up in Huber (2006).

Why are complete ranking functions confined to integers? The reason is condition (17.4). It entails that any infinite set of ranks has a minimum and hence that the range of a complete ranking function is well-ordered. Hence, the natural numbers are a natural choice. In my first publications (1983) and (1988) I allowed for more generality and assumed an arbitrary set of ordinal numbers as the range of a ranking function. However, since we want to calculate with ranks, this meant to engage into ordinal arithmetic, which is awkward. Therefore I later confined myself to complete ranking functions as defined above.

The issue about condition (17.4) was first raised by Lewis (1973, sect. 1.4) where he introduced the so-called Limit Assumption in relation to his semantics of counterfactuals. Endorsing (17.4), as I do, is tantamount to endorsing the Limit Assumption. Lewis finds reason against it, though it does not affect the *logic* of counterfactuals. From a semantic point of view, I do not understand his reason. He requests us to counterfactually suppose that a certain line is longer than an inch and asks how long it would or might be. He argues in effect that for each $\varepsilon > 0$ we should accept as true: "If the line would be longer than 1 inch, it would not be longer than $1 + \varepsilon$ inches." This strikes me as blatantly inconsistent, even if we cannot derive a contradiction in counterfactual logic (due to its $\omega$–incompleteness). Therefore, I am accepting the Limit Assumption and, correspondingly, the law of infinite disjunction. This means in particular that in that law the minimum must not be weakened to the infimum.

Though I prefer complete ranking functions for the reasons given, the issue will have no further relevance here. In particular, if we assume the algebra of propositions to be finite, each ranking function is complete, and the issue does not arise. In the sequel, you can add or delete completeness as you wish.

Let me add another observation apparently of a technical nature. It is that we can mix ranking functions in order to form a new ranking function. This is the content of

**Definition 17.3** Let $\Lambda$ be a non-empty set of negative ranking functions for an algebra $\mathcal{A}$ of propositions, and let $\rho$ be a complete negative ranking function over $\Lambda$. Then $\kappa$ defined by

$$\kappa(A) = \min\{\lambda(A) + \rho(\lambda) \mid \lambda \in \Lambda\} \text{ for all } A \in \mathcal{A} \qquad (17.5)$$

is obviously a negative ranking function for $\mathcal{A}$ as well and is called the *mixture of* $\Lambda$ *by* $\rho$.

It is nice that such mixtures make formal sense. However, we shall see in the course of this paper that the point is more than a technical one; such mixtures will acquire deep philosophical importance later on.

So far, (degree of) disbelief was our basic notion. Was this necessary? Certainly not. We might just as well express things in positive terms:

**Definition 17.4** Let $\mathcal{A}$ be an algebra over $W$. Then $\pi$ is a *positive ranking function* for $\mathcal{A}$ iff $\pi$ is a function from $\mathcal{A}$ into $\mathbf{R}^*$ such that for all $A, B \in \mathcal{A}$:

$$\pi(\varnothing) = 0 \text{ and } \pi(W) = \infty, \tag{17.6}$$

$$\pi(A \cap B) = \min\{\pi(A), \pi(B)\} \quad [\textit{the law of conjunction for positive ranks}]. \tag{17.7}$$

Positive ranks express *degrees of belief*. $\pi(A) > 0$ says that $A$ is believed (to some positive degree), and $\pi(A) = 0$ says that $A$ is not believed. Obviously, positive ranks are the dual to negative ranks; if $\pi(A) = \kappa(\overline{A})$ for all $A \in \mathcal{A}$, then $\pi$ is a positive function iff $\kappa$ is a negative ranking function.

Positive ranking functions seem distinctly more natural. Why do I still prefer the negative version? A superficial reason is that we have seen complete negative ranking functions to be reducible to point functions, whereas it would obviously be ill-conceived to try the same for the positive version. This, however, is only indicative of the main reason. Despite appearances, we shall soon see that negative ranks behave very much like probabilities. In fact, this parallel will serve as our compass for a host of exciting observations. (For instance, in the finite case probability measures can also be reduced to point functions.) If we were thinking in positive terms, this parallel would remain concealed.

There is a further notion that may appear even more natural:

**Definition 17.5** Let $\mathcal{A}$ be an algebra over $W$. Then $\tau$ is a *two-sided ranking function*[7] for $\mathcal{A}$ iff $\tau$ is a function from $\mathcal{A}$ into $\mathbf{R} \cup \{-\infty, \infty\}$ such that there is a negative ranking function $\kappa$ and its positive counterpart $\pi$ for which for all $A \in \mathcal{A}$:

$$\tau(A) = \kappa(\overline{A}) - \kappa(A) = \pi(A) - \kappa(A).$$

Obviously, we have $\tau(A) > 0$, $< 0$, or $= 0$ according to whether $A$ is believed, disbelieved, or neither. In this way, the belief values of all propositions are expressed in a single function. Moreover, we have the appealing law that $\tau(\overline{A}) = -\tau(A)$. For some purposes this is a useful notion that I shall readily employ. However, its formal behavior is awkward. Its direct axiomatic characterization would have been cumbersome, and its simplest definition consisted in its reduction to the other notions.

[7] In earlier papers I called this a belief function, obviously an unhappy term which has too many different uses. This is one reason fort the mild terminological reform proposed in this paper.

Still, this notion suggests an interpretational degree of freedom so far unnoticed.[8] We might ask: Why does the range of belief extend over all the positive reals in a two-sided ranking function and the range of disbelief over all the negative reals, whereas neutrality shrinks to rank 0? This looks unfair. Why may unopinionatedness not occupy a much broader range? Indeed, why not? We might just as well distinguish some positive rank or real $z$ and define the closed interval $[-z, z]$ as the range of neutrality. Then $\tau(A) > z$ expresses belief in $A$ and $\tau(A) < -z$ disbelief in $A$. This is a viable interpretation; in particular, consistency and deductive closure of belief sets would be preserved. However, 0 would still be a distinguished rank in this interpretation; it marks *central* neutrality, as it were, since it is the only rank $x$ for which we may have $\tau(A) = \tau(\overline{A}) = x$.

The interpretational freedom appears quite natural. After all, the notion of belief is certainly vague and can be taken more or less strict. We can do justice to this vagueness with the help of the parameter $z$. The crucial point, though, is that we always get the formal structure of belief we want to get, however we fix that parameter. The principal lesson of this observation is, hence, that it is not the notion of belief which is of basic importance; it is rather the formal structure of ranks. The study of belief *is* the study of *that* structure. Still, it would be fatal to simply give up talking of belief in favor of ranks. Ranks express beliefs, even if there is interpretational freedom. Hence, it is of paramount importance to maintain the intuitive connection. In the sequel, I shall stick to my standard interpretation and equate belief in $A$ with $\tau(A) > 0$, even though this is a matter of decision.

Let us pause for a moment and take a brief look back. What I have told so far probably sounds familiar. One has quite often seen all this, in this or a similar form – where the similar form may also be a comparative one: as long as only the ordering and not the numerical properties of the degrees of belief are relevant, a ranking function may also be interpreted as a weak ordering of propositions according to their plausibility, entrenchment, credibility, etc. Often things are cast in negative terms, as I primarily do, and often in positive terms. In particular, the law of negation securing consistency and the law of disjunction somehow generalizing deductive closure (we still have to look at the point more thoroughly) or their positive counterparts are pervasive. If one wants to distinguish a common core in that ill-defined family of Baconian probability, it is perhaps just these two laws.

So, why invent a new name, 'ranks', for familiar stuff? The reason lies in the second fundamental aim associated with ranking functions: to account for the dynamics of belief. This aim has been little pursued under the label of Baconian probability, but it is our central topic for the rest of this section. Indeed, everything stands and falls with our notion of conditional ranks; it is the distinctive mark of ranking theory. Here it is:

**Definition 17.6** Let $\kappa$ be a negative ranking function for $\mathcal{A}$ and $\kappa(A) < \infty$. Then the *conditional rank* of $B \in \mathcal{A}$ *given* $A$ is defined as $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$.

[8] I am grateful to Matthias Hild for making this point clear to me.

The function $\kappa_A : B \mapsto \kappa(B \mid A))$ is obviously a negative ranking function in turn and called the *conditionalization of* $\kappa$ *by* $A$.

We might rewrite this definition as a law:

$$\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A) \quad [\textit{the law of conjunction (for negative ranks)}]. \tag{17.8}$$

This amounts to the highly intuitive assertion that one has to add the degree of disbelief in $B$ given $A$ to the degree of disbelief in $A$ in order to get the degree of disbelief in $A$-and-$B$.

Moreover, it immediately follows for all $A, B \in \mathcal{A}$ with $\kappa(A) < \infty$:

$$\kappa(B \mid A) = 0 \text{ or } \kappa(\overline{B} \mid A) = 0 \quad [\textit{conditional law of negation}]. \tag{17.9}$$

This law says that even conditional belief must be consistent. If both, $\kappa(B \mid A)$ and $\kappa(\overline{B} \mid A)$, were $> 0$, both, $B$ and $\overline{B}$, would be believed given $A$, and this ought to be excluded, as long as the condition $A$ itself is considered possible.

Indeed, my favorite axiomatization of ranking theory runs reversely, it consists of the definition of conditional ranks and the conditional law of negation. The latter says that min $\{\kappa(A \mid A \cup B), \kappa(B \mid A \cup B)\} = 0$, and this and the definition of conditional ranks entail that min $\{\kappa(A), \kappa(B)\} = \kappa(A \cup B)$, i.e., the law of disjunction. Hence, the only substantial assumption written into ranking functions is conditional consistency, and it is interesting to see that this entails deductive closure as well. Huber (2007) has further improved upon this important idea and shown that ranking theory is indeed nothing but the assumption of dynamic consistency, i.e., the preservation of consistency under any dynamics of belief. (He parallels, in a way, the dynamic Dutch book argument for probabilities by replacing its assumption of no sure loss by the assumption of consistency under all circumstances.)

It is instructive to look at the positive counterpart of negative conditional ranks. If $\pi$ is the positive ranking function corresponding to the negative ranking function $\kappa$, Definition 17.6 simply translates into: $\pi(B \mid A) = \pi(\overline{A} \cup B) - \pi(\overline{A})$. Defining $A \to B = \overline{A} \cup B$ as set-theoretical 'material implication', we may as well write:

$$\pi(A \to B) = \pi(B \mid A) + \pi(\overline{A}) \quad [\textit{the law of material implication}]. \tag{17.10}$$

Again, this is highly intuitive. It says that the degree of belief in the material implication $A \to B$ is added up from the degree of belief in its vacuous truth (i.e., in $\overline{A}$) and the conditional degree of belief of $B$ given $A$.[9] However, again comparing the negative and the positive version, one can already sense the analogy between probability and ranking theory from (17.8), but hardly from (17.10). This analogy will play a great role in the following subsections.

[9]Thanks again to Matthias Hild for pointing this out to me.

Two-sided ranks have a conditional version as well; it is straightforward. If $\tau$ is the two-sided ranking function corresponding to the negative $\kappa$ and the positive $\pi$, then we may simply define:

$$\tau(B \mid A) = \pi(B \mid A) - \kappa(B \mid A) = \kappa(\overline{B} \mid A) - \kappa(B \mid A). \tag{17.11}$$

It will sometimes be useful to refer to these two-sided conditional ranks.

For illustration of negative conditional ranks, let us briefly return to our example of Tweetie. Above, I already mentioned various examples of if-then sentences, some held vacuously true and some non-vacuously. Now we can see that precisely the if-then sentences non-vacuously held true correspond to conditional beliefs. According to the $\kappa$ specified, you believe, e.g., that Tweetie can fly given it is a bird (since $\kappa(\overline{F} \mid B) = 1$)) and also given it is a bird, but not a penguin (since $\kappa(\overline{F} \mid B\&\overline{P}) = 2$)), that Tweetie cannot fly given it is a penguin (since $\kappa(F \mid P) = 3$) and even given it is a penguin, but not a bird (since $\kappa(F \mid \overline{B}\&P) = 3$)). You also believe that it is not a penguin given it is a bird (since $\kappa(P \mid B) = 1$) and that it is a bird given it is a penguin (since $\kappa(\overline{B} \mid P) = 7$)). And so forth.

Let us now unfold the power of conditional ranks and their relevance to the dynamics of belief in several steps.

## *Reasons and Their Balance*

The first application of conditional ranks is in the theory of confirmation. Basically, Carnap (1950) told us, confirmation is positive relevance. This idea can be explored probabilistically, as Carnap did. But here the idea works just as well. A proposition $A$ confirms or supports or speaks for a proposition $B$, or, as I prefer to say, $A$ is a reason for $B$, if $A$ strengthens the belief in $B$, i.e., if $B$ is more strongly believed given $A$ than given $\overline{A}$, i.e., iff $A$ is positively relevant for $B$. This is easily translated into ranking terms:

**Definition 17.7** Let $\kappa$ be a negative ranking function for $\mathcal{A}$ and $\tau$ the associated two-sided ranking function. Then $A \in \mathcal{A}$ is a *reason for* $B \in \mathcal{A}$ w.r.t. $\kappa$ iff $\tau(B \mid A) > \tau(B \mid \overline{A})$, i.e., iff $\kappa(\overline{B} \mid A) > \kappa(\overline{B} \mid \overline{A})$ or $\kappa(B \mid A) < \kappa(B \mid \overline{A})$.

If $P$ is a standard probability measure on $\mathcal{A}$, then probabilistic positive relevance can be expressed by $P(B \mid A) > P(B)$ or by $P(B \mid A) > P(B \mid \overline{A})$. As long as all three terms involved are defined, the two inequalities are equivalent. Usually, then, the first inequality is preferred because its terms may be defined while not all terms of the second inequality are defined. If $P$ is a Popper measure, this argument does not hold, and then it is easily seen that the second inequality is more adequate, just as in the case of ranking functions.[10]

[10] A case in point is the so-called problem of old evidence, which has a simple solution in terms of Popper measures and the second inequality; cf. Joyce (1999, pp. 203ff.).

Confirmation or support may take four different forms relative to ranking functions, which are unfolded in

**Definition 17.8** Let κ be a negative ranking function for $\mathcal{A}$, τ the associated two-sided ranking function, and $A, B \in \mathcal{A}$. Then

$$A \text{ is a } \left\{ \begin{array}{c} \textit{additional} \\ \textit{sufficient} \\ \textit{necessary} \\ \textit{insufficient} \end{array} \right\} \textit{reason for } B \text{ w.r.t. } \kappa \text{ iff } \left\{ \begin{array}{c} \tau(B \mid A) > \tau(B \mid \overline{A}) > 0 \\ \tau(B \mid A) > 0 \geq \tau(B \mid \overline{A}) \\ \tau(B \mid A) \geq 0 > \tau(B \mid \overline{A}) \\ 0 > \tau(B \mid A) > \tau(B \mid \overline{A}) \end{array} \right\}.$$

If $A$ is a reason for $B$, it must obviously take one of these four forms; and the only way to have two forms at once is by being a necessary and sufficient reason.[11]

Talking of reasons here is, I find, natural, but it stirs a nest of vipers. There is a host of philosophical literature pondering about reasons, justifications, etc. Of course, this is a field where multifarious philosophical conceptions clash, and it is not easy to gain an overview over the fighting parties. Here is not the place for starting a philosophical argument[12], but by using the term 'reason' I want at least to submit the claim that the topic may gain enormously by giving a central place to the above explication of reasons.

To elaborate only a little bit: When philosophers feel forced to make precise their notion of a (theoretical, not practical) reason, they usually refer to the notion of a *deductive* reason, as fully investigated in deductive logic. The deductive reason relation is reflexive, transitive, and not symmetric. By contrast, Definition 17.7 captures the notion of a *deductive or inductive* reason. The relation embraces the deductive relation, but it is reflexive, symmetric, and not transitive. Moreover, the fact that reasons may be additional or insufficient reasons according to Definition 17.8 has been neglected by the relevant discussion, which was rather occupied with necessary and/or sufficient reasons. Pursue, though, the use of the latter terms throughout the history of philosophy. Their deductive explication is standard and almost always fits. Often, it is clear that the novel inductive explication given by Definition 17.8 would be inappropriate. Very often, however, the texts are open to that inductive explication as well, and systematically trying to reinterpret these old texts would yield a highly interesting research program in my view.

The topic is obviously inexhaustible. Let me take up only one further aspect. Intuitively, we weigh reasons. This is a most important activity of our mind. We do not only weigh practical reasons in order to find out what to do, we also weigh theoretical reasons. We are wondering whether or not we should believe $B$, we are searching for reasons speaking in favor or against $B$, we are weighing these reasons, and we hopefully reach a conclusion. I am certainly not denying the phenomenon of

[11]In earlier publications I spoke of weak instead of insufficient reasons. Thanks to Arthur Merin who suggested the more appropriate term to me.

[12]I attempted to give a partial overview and argument in Spohn (2001a).

inference that is also important, but what is represented as an inference often rather takes the form of such a weighing procedure. 'Reflective equilibrium' is a familiar and somewhat more pompous metaphor for the same thing.

If the balance of reasons is such a central phenomenon the question arises: how can epistemological theories account for it? The question is less well addressed than one should think. However, the fact that there is a perfectly natural Bayesian answer is a very strong and more or less explicit argument in favor of Bayesianism. Let us take a brief look at how that answer goes:

Let $P$ be a (subjective) probability measure over $\mathcal{A}$ and let $B$ be the focal proposition. Let us look at the simplest case, consisting of one reason $A$ for $B$ and the automatic counter-reason $\overline{A}$ against $B$. Thus, in analogy to Definition 17.7, $P(B \mid A) > P(B \mid \overline{A})$. How does $P$ balance these reasons and thus fit in $B$? The answer is simple, we have:

$$P(B) = P(B \mid A) \cdot P(A) + P(B \mid \overline{A}) \cdot P(\overline{A}) \,. \tag{17.12}$$

This means that the probabilistic balance of reason is a *beam balance* in the literal sense. The length of the lever is $P(B \mid A) - P(B \mid \overline{A})$; the two ends of the lever are loaded with the *weights* $P(A)$ and $P(\overline{A})$ of the reasons; $P(B)$ divides the lever into two parts of length $P(B \mid A) - P(B)$ and $P(B) - P(B \mid \overline{A})$ representing the *strength* of the reasons; and then $P(B)$ must be chosen so that the beam is in balance. Thus interpreted (17.12) is nothing but the law of levers.

Ranking theory has an answer, too, and I am wondering who else has. According to ranking theory, the balance of reasons works like a *spring balance*. Let $\kappa$ be a negative ranking function for $\mathcal{A}$, $\tau$ the corresponding two-sided ranking function, $B$ the focal proposition, and $A$ a reason for $B$. So, $\tau(B \mid A) > \tau(B \mid \overline{A})$. Again, it easily proved that always $\tau(B \mid A) \geq \tau(B) \geq \tau(B \mid \overline{A})$. But where in between is $\tau(B)$ located? A little calculation shows the following specification to be correct:

Let $x = \kappa(B \mid \overline{A}) - \kappa(B \mid A)$ and $y = \kappa(\overline{B} \mid A) - \kappa(\overline{B} \mid \overline{A})$. *Then*

(a) $x, y \geq 0$ and $\tau(B \mid A) - \tau(B \mid \overline{A}) = x + y$,

(b) $\tau(B) = \tau(B \mid \overline{A})$, if $\tau(A) \leq -x$, (17.13)

(c) $\tau(B) = \tau(B \mid A)$, if $\tau(A) \geq y$,

(d) $\tau(B) = \tau(A) + \tau(B \mid \overline{A}) + x$, if $-x < \tau(A) < y$.

This does not look as straightforward as the probabilistic beam balance. Still, it is not so complicated to interpret (17.13) as a spring balance. The idea is that you hook in the spring at a certain point, that you extend it by the force of reasons, and that $\tau(B)$ is where the spring extends. Consider first the case where $x, y > 0$. Then you hook in the spring at point $0 \; (= \tau(B \mid \overline{A}) + x)$ and exert the force $\tau(A)$ on the spring. Either, this force transcends the lower stopping point $-x$ or the upper stopping point $y$. Then the spring extends exactly till the stopping point, as (17.13b + c) say. Or, the force $\tau(A)$ is less. Then the spring extends exactly by $\tau(A)$, according to (17.13d).

The second case is that $x = 0$ and $y > 0$. Then you fix the spring at $\tau(B \mid \overline{A})$, the lower point of the interval in which $\tau(B)$ can move. The spring cannot extend below that point, says (17.13b). But according to (17.13c + d) it can extend above, by the force $\tau(A)$, but not beyond the upper stopping point. For the third case $x > 0$ and $y = 0$ just reverse the second picture. In this way, the force of the reason $A$, represented by its two-sided rank $\tau(A)$, pulls the two-sided rank of the focal proposition $B$ to its proper place within the interval $[\tau(B \mid \overline{A}), \tau(B \mid A)]$ fixed by the relevant conditional ranks.

I do not want to assess these findings in detail. You might prefer the probabilistic balance of reasons, a preference I would understand. You might be happy to have at least one alternative model, an attitude I recommend. Or you may search for further models of the weighing of reasons; in this case, I wish you good luck. What you may not do is ignoring the issue; your epistemology is incomplete if it does not take a stance. And one must be clear about what is required for taking a stance. As long as one considers positive relevance to be the basic characteristic of reasons, one must provide some notion of conditional degrees of belief, conditional probabilities, conditional ranks, or whatever. Without some well-behaved conditionalization one cannot succeed.

## *The Dynamics of Belief and the Measurement of Belief*

Our next point will be to define a reasonable dynamics for ranking functions that entails a dynamic for belief. There are many causes which affect our beliefs, forgetfulness as a necessary evil, drugs as an unnecessary evil, and so on. From a rational point of view, it is scarcely possible to say anything about such changes.[13] The rational changes are due to experience or information. Thus, it seems we have already solved our task: if $\kappa$ is my present doxastic state and I get informed about the proposition $A$, then I move to the conditionalization $\kappa_A$ of $\kappa$ by $A$. This, however, would be a bad idea. Recall that we have $\kappa_A(\overline{A}) = \infty$, i.e., $A$ is believed with absolute certainty in $\kappa_A$; no future evidence could cast any doubt on the information. This may sometimes happen; but usually information does not come so firmly. Information may turn out wrong, evidence may be misleading, perception may be misinterpreted; we should provide for flexibility. How?

One point of our first attempt was correct; if my information consists solely in the proposition $A$, this cannot affect my beliefs conditional on $A$. Likewise it cannot affect my beliefs conditional on $\overline{A}$. Thus, it directly affects only how firmly I believe $A$ itself. So, how firmly should I believe $A$? There is no general answer. I propose to turn this into a parameter of the information process itself; somehow the way I get informed about $A$ entrenches $A$ in my belief state with a certain firmness $x$.

[13] Although there is a (by far not trivial) decision rule telling that costless memory is never bad, just as costless information; cf. Spohn (1976/78, sect. 4.4).

The point is that as soon as the parameter is fixed and the constancy of the relevant conditional beliefs is accepted, my posterior belief state is fully determined. This is the content of

**Definition 17.9** Let $\kappa$ be a negative ranking function for $\mathcal{A}$, $A \in \mathcal{A}$ such that $\kappa(A)$, $\kappa(\overline{A}) < \infty$, and $x \in \mathbf{R}^*$. Then the $A \to x$-conditionalization $\kappa_{A \to x}$ of $\kappa$ is defined by $\kappa_{A\to x}(B) = \begin{cases} \kappa(B \mid A) & \text{for } B \subseteq A, \\ \kappa(B \mid \overline{A}) + x & \text{for } B \subseteq \overline{A} \end{cases}$. From this $\kappa_{A\to x}(B)$ may be inferred for all other $B \in \mathcal{A}$ with the law of disjunction.

Hence, the effect of the $A \to x$-conditionalization is to shift the possibilities in $A$ (to lower ranks) so that $\kappa_{A\to x}(A) = 0$ and the possibilities in $\overline{A}$ (to higher ranks) so that $\kappa_{A\to x}(\overline{A}) = x$. If one is attached to the idea that evidence consists in nothing but a proposition, the additional parameter is a mystery. The processing of evidence may indeed be so automatic that one hardly becomes aware of this parameter. Still, I find it entirely natural that evidence comes more or less firmly. Consider, for instance, the proposition: "There are tigers in the Amazon jungle", and consider six scenarios: (a) I read a somewhat sensationalist coverage in the yellow press claiming this, (b) I read a serious article in a serious newspaper claiming this, (c) I hear the Brazilian government officially announcing that tigers have been discovered in the Amazon area, (d) I see a documentary in TV claiming to show tigers in the Amazon jungle, (e) I read an article in *Nature* by a famous zoologist reporting of tigers there, (f) I travel there by myself and see the tigers. In all six cases I receive the information that there are tigers in the Amazon jungle, but with varying and, I find, increasing certainty.

One might object that the evidence and thus the proposition received is clearly a different one in each of the scenarios. The crucial point, though, is that we are dealing here with a fixed algebra $\mathcal{A}$ of propositions and that we have nowhere presupposed that this algebra consists of all propositions whatsoever; indeed, that would be a doubtful presupposition. Hence $\mathcal{A}$ may be course-grained and unable to represent the propositional differences between the scenarios; the proposition in $\mathcal{A}$ which is directly affected in the various scenarios may be just the proposition that there are tigers in the Amazon jungle. Still the scenarios may be distinguished by the firmness parameter.

So, the dynamics of ranking functions I propose is simply this: Suppose $\kappa$ is your prior doxastic state. Now you receive some information $A$ with firmness $x$. Then your posterior state is $\kappa_{A\to x}$. Your beliefs change accordingly; they are what they are according to $\kappa_{A\to x}$. Note that the procedure is iterable. Next, you receive the information $B$ with firmness $y$, and so you move to $(\kappa_{A\to x})_{B\to y}$. And so on. This point will acquire great importance later on.

I should mention, though, that this iterability need not work in full generality. Let us call a negative ranking function $\kappa$ *regular* iff $\kappa(A) < \infty$ for all $A \neq \varnothing$. Then we obviously have that $\kappa_{A\to x}$ is regular if $\kappa$ is regular and $x < \infty$. Within the realm of regular ranking functions iteration of changes works without restriction. Outside this realm you may get problems with the rank $\infty$.

There is an important generalization of Definition 17.9. I just made a point of the fact that the algebra $\mathcal{A}$ may be too coarse-grained to propositionally represent all possible evidence. Why assume then that it is just one proposition $A$ in the algebra that is directly affected by the evidence? Well, we need not assume this. We may more generally assume that the evidence affects some evidential partition $\mathcal{E}$ of $W$ and assigns some new ranks to the members of the partition, which we may sum up in a complete ranking function $\lambda$ on $\mathcal{E}$. Then we may define the $\mathcal{E} \rightarrow \lambda$-*conditionalization* $\kappa_{\mathcal{E}\rightarrow\lambda}$ of the prior $\kappa$ by $\kappa_{\mathcal{E}\rightarrow\lambda}(B) = \kappa(B \mid E_i) + \lambda(E_i)$ for $B \subseteq E_i$ $(i = 1, \ldots, n)$ and infer $\kappa_{\mathcal{E}\rightarrow\lambda}(B)$ for all other $B$ by the law of disjunction. This is the most general law of doxastic change in terms of ranking functions I can conceive of. Note that we may describe the $\mathcal{E} \rightarrow \lambda$-conditionalization of $\kappa$ as the mixture of all $\kappa_{E_i}$ $(i = 1, \ldots, n)$. So, this is a first useful application of mixtures of ranking functions.

Here, at last, the reader will have noticed the great similarity of my conditionalization rules with Jeffrey's probabilistic conditionalization first presented in Jeffrey (1965, ch. 11). Indeed, I have completely borrowed my rules from Jeffrey. Still, let us further defer the comparison of ranking with probability theory. The fact that many things run similarly does not mean that one can dispense with the one in favor of the other, as I shall make clear in section "Ranks and probabilities", pp. 328ff.

There is an important variant of Definition 17.9. Shenoy (1991), and several authors after him, pointed out that the parameter $x$ as conceived in Definition 17.9 does not characterize the evidence as such, but rather the result of the interaction between the prior doxastic state and the evidence. Shenoy proposed a reformulation with a parameter exclusively pertaining to the evidence:

**Definition 17.10** Let $\kappa$ be a negative ranking function for $\mathcal{A}$, $A \in \mathcal{A}$ such that $\kappa(A)$, $\kappa(\overline{A}) < \infty$, and $x \in \mathbf{R}^*$. Then the $A\uparrow x$-*conditionalization* $\kappa_{A\uparrow x}$ of $\kappa$ is defined by $\kappa_{A\uparrow x}(B) = \begin{cases} \kappa(B) - y \text{ for } B \subseteq A, \\ \kappa(B) + x - y \text{ for } B \subseteq \overline{A}, \end{cases}$ where $y = \min\{\kappa(A), x\}$. Again, $\kappa_{A\uparrow x}(B)$ may be inferred for all other $B \in \mathcal{A}$ by the law of disjunction.

The effect of this conditionalization is easily stated. It is, whatever the prior ranks of $A$ and $\overline{A}$ are, that the possibilities within $A$ improve by exactly $x$ ranks in comparison to the possibilities within $\overline{A}$. In other words, we always have $\tau_{A\uparrow x}(A) - \tau(A) = x$ (in terms of the prior and the posterior two-sided ranking function).

It is thus appropriate to say that in $A\uparrow x$-conditionalization the parameter $x$ exclusively characterizes the evidential impact. We may characterize the $A\rightarrow x$-conditionalization of Definition 17.9 as *result-oriented* and the $A\uparrow x$-conditionalization of Definition 17.10 as *evidence-oriented*. Of course, the two variants are easily interdefinable. We always have $\kappa_{A\rightarrow x} = \kappa_{A\uparrow y}$, where $y = x - \tau(A) = x + \tau(\overline{A})$. Still, it is sometimes useful to change perspective from one variant to the other.[14]

[14]Generalized probabilistic conditionalization as originally proposed by Jeffrey was result-oriented as well. However, Garber (1980) observed that there is also an evidence-oriented version of generalized probabilistic conditionalization. The relation, though, is not quite as elegant.

For instance, the evidence-oriented version helps to some nice observations. We may note that conditionalization is reversible: $(\kappa_{A\uparrow x})_{\overline{A}\uparrow x} = \kappa$. So, there is always a possible second change undoing the first. Moreover, changes always commute: $(\kappa_{A\uparrow x})_{B\uparrow_y} = (\kappa_{B\uparrow y})_{A\uparrow x}$. In terms of result-oriented conditionalization this law would look more awkward. Commutativity does not mean, however, that one could comprise the two changes into a single change. Rather, the joint effect of two conditionalizations according to Definition 17.9 or 17.10 can in general only be summarized as one step of generalized $\mathcal{E} \to \lambda$-conditionalization. I think that reversibility and commutativity are intuitively desirable.

Change through conditionalization is driven by information, evidence, or perception. This is how I have explained it. However, we may also draw a more philosophical picture, we may also say that belief change according to Definition 17.9 or 17.10 is driven by reasons. Propositions for which the information received is irrelevant do not change their ranks, but propositions for which that information is positively or negatively relevant do change their ranks. The evidential force pulls at the springs and they must find a new rest position for all the propositions for or against which the evidence speaks, just in the way I have described in the previous subsection.

This is a strong picture captivating many philosophers. However, I have implemented it in a slightly unusual way. The usual way would have been to attempt to give some substantial account of what reasons are on which an account of belief dynamics is thereafter based. I have reversed the order. I have first defined conditionalization in Definition 17.6 and the more sophisticated form in Definitions 17.9 and 17.10. With the help of conditionalization, i.e., from this account of belief dynamics, I could define the reason relation in a way sustaining this picture. At the same time this procedure entails dispensing with a more objective notion of a reason. Rather, what is a reason for what is entirely determined by the subjective doxastic state as represented by the ranking function at hand. Ultimately, this move is urged by inductive skepticism as enforced by David Hume and reinforced by Nelson Goodman. But it does not mean surrender to skepticism. On the contrary, we are about to unfold a positive theory of rational belief and rational belief change, and we shall have to see how far it carries us.[15]

If one looks at the huge literature on belief change, one finds discussed predominantly three kinds of changes: expansions, revisions, and contractions. Opinions widely diverge concerning these three kinds. For Levi, for instance, revisions are whatever results form concatenating contractions and expansions according to the so-called Levi identity, and so he investigates the latter (see his most recent account in Levi 2004). The AGM approach characterizes both, revisions and contractions, and claims nice correspondences back and forth by help of the Levi and the Harper identity (cf., e.g., Gärdenfors 1988, chs. 3 and 4). Or one might object to the characterization of contraction, but accept that of revision, and hence reject these identities. And so forth.

---

[15]Here it does not carry us far beyond the beginnings. In Spohn (1991, 1999) I have argued for some stronger rationality requirements and their consequences.

I do not really want to discuss the issue. I only want to point out that we have already taken a stance insofar as expansions, revisions, and contractions are all special cases of our $A \to x$–conditionalization. This is more easily explained in terms of result-oriented conditionalization:

If $\kappa(A) = 0$, i.e., if $A$ is not disbelieved, then $\kappa_{A\to x}$ represents an *expansion* by $A$ for any $x > 0$. If $\kappa(\overline{A}) = 0$, the expansion is genuine, if $\kappa(\overline{A}) > 0$, i.e., if $A$ is already believed in $\kappa$, the expansion is vacuous. Are there many different expansions? Yes and no. Of course, for each $x > 0$ a different $\kappa_{A\to x}$ results. On the other hand, one and the same belief set is associated with all these expansions. Hence, the expanded belief set is uniquely determined.

Similarly for revision. If $\kappa(A) > 0$, i.e., if $A$ is disbelieved, then $\kappa_{A\to x}$ represents a genuine *revision* by $A$ for any $x > 0$. In this case, the belief in $\overline{A}$ must be given up and along with it many other beliefs; instead, $A$ must be adopted together with many other beliefs. Again, there are many different revisions, but all of them result in the same revised belief set.

Finally, if $\kappa(A) = 0$, i.e., if $A$ is not disbelieved, then $\kappa_{A\to 0}$ represents contraction by $A$. If $\kappa(\overline{A}) > 0$, i.e., if $A$ is even believed, the contraction is genuine; then belief in $A$ is given up after contraction and no new belief adopted. If $\kappa(\overline{A}) = 0$, the contraction is vacuous; there was nothing to contract in the first place. If $\kappa(A) > 0$, i.e., if $\overline{A}$ is believed, then $\kappa_{A\to 0} = \kappa_{\overline{A}\to 0}$ rather represents contraction by $\overline{A}$.[16]

As observed in Spohn (1988, footnote 20) and more fully explained in Gärdenfors (1988, pp. 73f.), it is easily checked that expansions, revisions, and contractions thus defined satisfy all of the original AGM postulates (K*1-8) and (K$^-$1-8) (cf. Gärdenfors 1988, pp. 54–56 and 61–64) (when they are translated from AGM's sentential framework into our propositional or set-theoretical one). For those like me who accept the AGM postulates this is a welcome result.

For the moment, though, it may seem that we have simply reformulated AGM belief revision theory. This is not so; $A \to x$-conditionalization is much more general than the three AGM changes. This is clear from the fact that there are many different expansions and revisions that cannot be distinguished by the AGM account. It is perhaps clearest in the case of vacuous expansion that is no change at all in the AGM framework, but may well be a genuine change in the ranking framework, a redistribution of ranks which does not affect the surface of beliefs. Another way to state the same point is that insufficient and additional reasons also drive doxastic changes, which, however, are inexpressible in the AGM framework. For instance, if $A$ is still disbelieved in the $A\uparrow x$-*conditionalization* $\kappa_{A\uparrow x}$ of $\kappa$ (since $\kappa(A) > x$), one has obviously received only an insufficient reason for $A$, and the $A\uparrow x$-*conditionalization* might thus be taken to represent what is called non-prioritized belief revision in the AGM literature (cf. Hansson 1997).

[16] If we accept the idea in section "Basics" (p. 311) of taking the interval $[-z, z]$ of two-sided ranks as the range of neutrality, contraction seems to become ambiguous as well. However, the contraction just defined would still be distinguishable as a *central* contraction since it gives the contracted proposition central neutrality.

This is not the core of the matter, though. The core of the matter is *iterated belief change*, which I have put into the center of my considerations in Spohn (1983, sect. 5.3, 1988). As I have argued there, AGM belief revision theory is essentially unable to account for iterated belief change. I take 20 years of multifarious, but in my view unsatisfactory attempts to deal with that problem (see the overview in Rott 2008) as confirming my early assessment. By contrast, changes of the type $A \to x$-conditionalization are obviously indefinitely iterable.

In fact, my argument in Spohn (1988) was stronger. It was that if AGM belief revision theory is to be improved so as to adequately deal with the problem of iterated belief change, ranking theory is the only way to do it. I always considered this to be a conclusive argument in favor of ranking theory.

This may be so. Still, AGM theorists, and others as well, remained skeptical. "What exactly is the meaning of numerical ranks?" they asked. One may well acknowledge that the ranking apparatus works in a smooth and elegant way, has a lot of explanatory power, etc. But all this does not answer this question. Bayesians have met this challenge. They have told stories about the operational meaning of subjective probabilities in terms of betting behavior, they have proposed an ingenious variety of procedures for measuring this kind of degrees of belief. One would like to see a comparative achievement for ranking theory.

It exists and is finally presented in Hild and Spohn (2008). There is no space here to fully develop the argument. However, the basic point can easily be indicated so as to make the full argument at least plausible. The point is that ranks do not only account for iterated belief change, but can reversely be measured thereby. This may at first sound unhelpful. $A \to x$-conditionalization refers to the number $x$; so even if ranks can somehow be measured with the help of such conditionalizations, we do not seem to provide a fundamental measurement of ranks. Recall, however, that (central) contraction by $A$ (or $\overline{A}$) is just $A \to 0$-conditionalization and is thus free of a hidden reference to numerical ranks; it only refers to rank 0 which has a clear operational or surface interpretation in terms of belief. Hence, the idea is to measure ranks by means of iterated contractions; if that works, it really provides a fundamental measurement of ranks that is based only on the beliefs one now has and one would have after various iterated contractions.

How does the idea work? Recall our observation above that the positive rank of a material implication $A \to B$ is the sum of the degree of belief in $B$ given $A$ and the degree of belief in the vacuous truth of $A \to B$, i.e., of $\overline{A}$. Hence, after contraction by $\overline{A}$, belief in the material implication $A \to B$ is equivalent to belief in $B$ given $A$, i.e., to the positive relevance of $A$ to $B$. This is how the reason relation, i.e., positive relevance, manifests itself in beliefs surviving contractions. Similarly for negative relevance and irrelevance.

Next observe that positive relevance can be expressed by certain inequalities for ranks that compare certain differences between ranks (similarly for negative relevance and irrelevance). This calls for applying the theory of difference measurement, as paradigmatically presented by Krantz et al. (1971, ch. 4).

Let us illustrate how this might work in our Tweetie example, pp. 306f. There we had specified a ranking function $\kappa$ for the eight propositional atoms, entailing ranks for all 256 propositions involved. Focusing on the atoms, we are thus dealing with a realm $X = \{x_1, \ldots x_8\}$ (where $x_1 = B \,\&\, \overline{P} \,\&\, F$, etc.) and a numerical function $f$ such that

$$f(x_1) = 0,\ f(x_2) = 4,\ f(x_3) = 0,\ f(x_4) = 11,$$
$$f(x_5) = 2,\ f(x_6) = 1,\ f(x_7) = 0,\ f(x_8) = 8.$$

This induces a lot of difference comparisons. For instance, we have, $f(x_6) - f(x_5) < f(x_2) - f(x_1)$. It is easily checked that this inequality says that, given $B$ (being a bird), $P$ (being a penguin) is positively relevant to $\overline{F}$ (not being able to fly) and that this in turn is equivalent with $P \rightarrow \overline{F}$ or $\overline{P} \rightarrow F$ still being believed after iterated contraction first by $\overline{B}$ and then by $P$ and $\overline{P}$ (only one of the latter is a genuine contraction). Or we have $f(x_2) - f(x_6) = f(x_4) - f(x_8)$. Now, this is an equality saying that, given $P$, $B$ (and $\overline{B}$) is irrelevant to $F$ (and $\overline{F}$), and this in turn is equivalent with none of the four material implications from $B$ or $\overline{B}$ to $F$ or $\overline{F}$ being believed after iterated contraction first by $\overline{P}$ and then by $B$ and $\overline{B}$ (again, only one of the latter is a genuine contraction).

Do these comparisons help to determine $f$? Yes, the example was so constructed: First, we have $f(x_1) - f(x_3) = f(x_3) - f(x_1) = f(x_1) - f(x_7)$. This entails $f(x_1) = f(x_3) = f(x_7)$. Let us choose this as the zero point of our ranking scale; i.e., $f(x_1) = 0$. Next, we have $f(x_5) - f(x_6) = f(x_6) - f(x_1)$. If we choose $f(x_6) = 1$ as our ranking unit, this entails $f(x_5) = 2$. Then, we have $f(x_2) - f(x_5) = f(x_5) - f(x_1)$, entailing $f(x_2) = 4$, and $f(x_8) - f(x_2) = f(x_2) - f(x_1)$, entailing $f(x_8) = 8$. Finally, we have $f(x_4) - f(x_8) = f(x_2) - f(x_6)$, the equation I had already explained, so that $f(x_4) = 11$. In this way, the difference comparisons entailed by our specification of $f$ determine $f$ uniquely up to a unit and a zero point.

The theory of difference measurement tells us how this procedure works in full generality. The resulting theorem says the following: Iterated contractions behave thus and thus if and only if differences between ranks behave thus and thus; and if differences between ranks behave thus and thus, then there is a ranking function measured on a ratio scale, i.e., unique up to a multiplicative constant, which exactly represents these differences. (See theorems 4.12 and 6.21 in Hild and Spohn (2008) for what “thus and thus” precisely means.)

On the one hand, this provides for an axiomatization of iterated contraction going beyond Darwiche and Pearl (1997), who presented generally accepted postulates of iterated revision and contraction and partially agreeing and partially disagreeing with further postulates proposed.[17] This axiomatization is assessible on intuitive and other grounds. On the other hand, one knows that if one accepts this axiomatization of iterated contraction one is bound to accept ranks as I have proposed them. Ranks do not fall from the sky, then; on the contrary, they uniquely represent contraction behavior.

[17]For an overview over such proposals see Rott (2008). For somewhat more detailed comparative remarks see Hild and Spohn (2008, sect. 5).

## *Conditional Independence and Bayesian Nets*

It is worthwhile looking a bit more at the details of belief formation and revision. For this purpose we should give more structure to propositions. They have a Boolean structure so far, but we cannot yet compose them from basic propositions as we intuitively do. A common formal way to do this is to generate propositions from (random) variables. I identify a variable with the set of its possible values. I intend variables to be specific ones. E.g., the temperature at March 15, 2005, in Konstanz (not understood as the actual temperature, but as whatever it may be, say, between $-100$ and $+ 100$ °C) is such a variable. Or, to elaborate, if we consider each of the six general variables temperature, air pressure, wind, humidity, precipitation, cloudiness at each of the 500 weather stations in Germany twice a day at each of the 366 days of 2004, we get a collection of $6 \times 500 \times 732$ specific variables with which we can draw a detailed picture of the weather in Germany in 2004.

So, let $V$ be the set of specific variables considered, where each $v \in V$ is just at least a binary set. A possible course of events or a possibility, for short, is just a selection function $w$ for $V$, i.e., a function $w$ on $V$ such that $w(v) \in v$ for all $v \in V$. Hence, each such function specifies a way how the variables in $V$ may realize. The set of all possibilities then simply is $W = \times V$. As before, propositions are subsets of $W$. Now, however, we can say that propositions are *about* certain variables. Let $X \subseteq V$. Then we say that $w, w' \in W$ *agree on* $X$ iff $w(v) = w'(v)$ for all $v \in X$. And we define that a proposition $A$ is *about* $X \subseteq V$ iff, for each $w$ in $A$, all $w'$ agreeing with $w$ on $X$ are in $A$ as well. Let $\mathcal{A}(X)$ be the set of propositions about $X$. Clearly, $\mathcal{A}(X) \subseteq \mathcal{A}(Y)$ for $X \subseteq Y$, and $\mathcal{A} = \mathcal{A}(V)$. In this way, propositions are endowed with more structure. We may conceive of propositions about single variables as *basic* propositions; the whole algebra $\mathcal{A}$ is obviously generated by such basic propositions (at least if $V$ is finite). So much as preparation for the next substantial step.

This step consists in more closely attending to (doxastic) dependence and independence in ranking terms. In a way, we have already addressed this issue: dependence is just positive or negative relevance, and independence is irrelevance. Still, let me state

**Definition 17.11** Let $\kappa$ be a negative ranking function for $\mathcal{A}$ and $A, B, C \in \mathcal{A}$. Then $A$ and $B$ are *independent* w.r.t. $\kappa$, i.e., $A \perp B$, iff $\tau(B \mid A) = \tau(B \mid \overline{A})$, i.e., iff for all $A' \in \{A, \overline{A}\}$ and $B' \in \{B, \overline{B}\}$ $\kappa(A' \cap B') = \kappa(A') + \kappa(B')$. And $A$ and $B$ are *independent given* $C$ w.r.t. $\kappa$, i.e., $A \perp B \,/\, C$, iff $A$ and $B$ are independent w.r.t. $\kappa_C$.

(Conditional) independence is symmetric. If $A$ is independent from $B$, $\overline{A}$ is so as well. If $A$ is independent from $B$ and $A'$ disjoint from $A$, then $A'$ is independent from $B$ iff $A \cup A'$ is. $\emptyset$ and $W$ are independent from all propositions. And so on.

The more interesting notion, however, is dependence and independence among variables. Look at probability theory where research traditionally and overwhelmingly focused on independent series of random variables and on Markov processes that are characterized by the assumption that past and future variables are independent given the present variable. We have already prepared for explaining this notion in ranking terms as well.

**Definition 17.12** Let $\kappa$ be a ranking function for $\mathcal{A} = \mathcal{A}(V)$, and let $X, Y, Z \subseteq V$ be sets of variables. Then $X$ and $Y$ are *independent* w.r.t. $\kappa$, i.e., $X \perp Y$, iff $A \perp B$ for all $A \in \mathcal{A}(X)$ and all $B \in \mathcal{A}(Y)$. Let moreover $\mathcal{Z}(Z)$ be the set of atoms of $\mathcal{A}(Z)$, i.e., the set of the logically strongest, non-empty proposition in $\mathcal{A}(Z)$. Then $X$ and $Y$ are *independent given* $Z$ w.r.t. $\kappa$, i.e., $X \perp Y / Z$, iff $A \perp B/C$ for all $A \in \mathcal{A}(X)$, $B \in \mathcal{A}(Y)$, and $C \in \mathcal{Z}(Z)$.

In other words, $X \perp Y/Z$ iff all propositions about $X$ are independent from all propositions about $Y$ given any full specification of the variables in $Z$. Conditional independence among sets of variables obey the following laws:

Let $\kappa$ be a negative ranking function for $\mathcal{A}(V)$. Then for any mutually disjoint $X, Y, Z, U \subseteq V$:

(a) if $X \perp Y/Z$, then $Y \perp X/Z$ [Symmetry],
(b) if $X \perp Y \cup U/Z$, then $X \perp Y/Z$ and $X \perp U/Z$ [Decomposition],
(c) $X \perp Y \cup U/Z$, then $X \perp Y/Z \cup U$ [Weak Union],
(d) $X \perp Y/Z$ and $X \perp U/Z \cup Y$, then $X \perp Y \cup U/Z$ [Contraction],
(e) if $\kappa$ is regular and if $X \perp Y/Z \cup U$ and $X \perp U/Z \cup Y$, then $X \perp Y \cup U/Z$ [Intersection]

(17.14)

These are nothing but what Pearl (1988, p. 88) calls the *graphoid* axioms; the labels are his (cf. p. 84). (Note that law (d), contraction, has nothing to do with contraction in belief revision theory.) That probabilistic conditional independence satisfies these laws was first proved in Spohn (1976/78, sect. 3.2) and Dawid (1979). The ranking Theorem (17.14) was proved in Spohn (1983, sect. 5.3, 1988, sect. 6). I conjectured in 1976, and Pearl conjectured, too, that the graphoid axioms give a complete characterization of conditional independence. We were disproved, however, by Studeny (1989) w.r.t. probability measures, but the proof carries over to ranking functions (cf. Spohn 1994a). Under special conditions, though, the graphoid axioms *are* complete, as was proved by Geiger and Pearl (1990) for probability measures and by Hunter (1991) for ranking functions (cf. again, Spohn 1994a).

I am emphasizing all this, because the main purport of Pearl's path-breaking book (1988) is to develop what he calls the theory of Bayesian nets, a theory that has acquired great importance and is presented in many text books (see, e.g., Neapolitan 1990 or Jensen 2001). Pearl makes very clear that the basis of this theory consists in the graphoid axioms; these allow representing conditional dependence and independence among sets of variables by Bayesian nets, i.e., by directed acyclic graphs, the nodes of which are variables. A vertex $u \rightarrow v$ of the graph then represents the fact that $v$ is dependent on $u$ given all the variables preceding $v$ in some given order, for instance, temporally preceding $v$. A major point of this theory is that it can describe in detail how probabilistic change triggered at some node in the net propagates throughout the net. All this is not merely mathematics, it is intuitively sensible and philosophically highly significant; for instance, inference acquires a novel and fruitful meaning in the theory of Bayesian nets.

Of course, my point now is that all these virtues carry over to ranking theory with the help of observation (17.14). The point is obvious, but hardly elaborated; that should be done.[18] It will thus turn out that ranks and hence beliefs can also be represented and computationally managed in that kind of structure.

This is not yet the end of the story. Spirtes et al. (1993) (see also Pearl 2000) have made amply clear that probabilistic Bayesian nets have a most natural causal interpretation; a vertex $u \rightarrow v$ then represents that the variable $v$ directly causally depends on the variable $u$. Spirtes et al. back up this interpretation, i.e., this connection of probability and causality, by their three basic axioms: the causal Markov condition, the minimality condition, and, less importantly, the faithfulness condition (cf. Spirtes et al. 1993, sect. 3.4). And they go on to develop a really impressive account of causation and causal inference on the basis of these axioms and thus upon the theory of Bayesian nets.

Again, all this carries over to ranking theory. Indeed, this is what ranks were designed for in the first place. In Spohn (1983) I gave an explication of probabilistic causation that entails the causal Markov condition and the minimality condition, and also Reichenbach's principle of the common cause, as I observed later in Spohn (1994b).[19] And I was convinced of the idea that, if the theory of causation is bound to bifurcate into a deterministic and a probabilistic branch, these two branches must at least be developed in perfect parallel. Hence, I proposed ranking theory in Spohn (1983) in order to realize this idea.[20] Of course, one has to discuss how adequate that theory of deterministic causation is, just as the adequacy of the causal interpretation of Bayesian nets is open to discussion. Here, my point is only that this deep philosophical perspective lies within reach of ranking theory; it is what originally drove that theory.

### *Objective Ranks?*

Now, a fundamental problem of ranking theory is coming into sight. I have emphasized that ranking functions represent rational beliefs and their rational dynamics and are thus entirely subject-bound. You have your ranking function and I have mine. We may or may not harmonize. In any case, they remain our subjective property.

---

[18] It has been done in the meantime. See Hohenadel (2013).

[19] I have analyzed the relation between Spirtes' et al. axiomatic approach to causation and my definitional approach a bit more thoroughly in Spohn (2001b).

[20] For a recent presentation of the account of deterministic causation in terms of ranking functions and its comparison in particular with David Lewis' counterfactual approach see Spohn (2006).

I have also emphasized the analogy to probability theory. There, however, we find subjective *and* objective probabilities. There are radicals who deny the one or the other kind of probability; and the nature of objective probabilities may still be ill understood. So, we certainly enter mined area here. Still, the predominant opinion is that both, the subjective and the objective notion, are somehow meaningful.

We therefore face a tension. It increases with our remarks about causation. I said I have provided an analysis of causation in ranking terms. If this analysis were to go through, the consequence would be that causal relations obtain relative to a ranking function, i.e., relative to the doxastic state of a subject. David Hume endorsed and denied this consequence at the same time; he was peculiarly ambiguous. This ambiguity must, however, be seen as his great achievement with which all philosophers after him had and still have to struggle. In any case, it will not do to turn causation simply into a subjective notion, as I seem to propose. If my strategy is to work at all, then the actually existing causal relations have to be those obtaining relative to the objectively correct ranking function. Is there any way to make sense of this phrase? (It is not even a notion yet.)

Yes, partially. The beginning is easy. Propositions are objectively true or false, and so are beliefs. Hence, a ranking function may be called objectively true or false as well, according to the beliefs it embodies. However, this is a very small step. Ranking functions can agree in their belief sets or in the propositions receiving rank 0, and yet widely diverge in the other ranks and thus in inductive and dynamic behavior. So, the suggested beginning is a very small step, indeed.

Taking a bigger step is more difficult. In my (1993) I have made a precise and detailed proposal that I still take to be sound; there is no space to repeat it here. Let me only briefly explain the basic idea. It is simply this: If propositions and beliefs are objectively true or false, then other features of ranking functions can be objectified to the extent to which these features are uniquely reflected in the associated belief sets. One constructive task is then to precisely define the content of the phrase 'uniquely reflected' and the required presuppositions or restrictions. The other constructive task is to inquire which specific features can in this sense be objectified to which specific extent.

Very roughly, the results in my (1993) are this: First, positive relevance, i.e., the reason relation, is *not* objectifiable in this sense, even if restricted to necessary and/or sufficient reasons. Second, whenever $A$ is a sufficient or necessary direct cause of $B$ w.r.t. $\kappa$, there is an associated material implication of the form "if the relevant circumstances obtain, then if $A$, then $B$, or, respectively, if $\overline{A}$, then $\overline{B}$". I call the conjunction of all these material implications the *causal law* associated with $\kappa$. The causal law is a proposition, an objective truth-condition. The point now is that there is a rich class of ranking functions which, under certain presuppositions, can uniquely be reconstructed from their causal laws and which may thus be called causal laws as well. In this sense and to this extent, causal relations obtaining relative to a subjective raking function can be objectified and thus do hold objectively.

A special case treated in Spohn (2002, 2005a) is the case of strict or deterministic laws. A strict law is, by all means, a regularity, an invariable obtaining of a certain type of state of affairs. But not any regularity is a law. What I have proposed in

Spohn (2002) is that a law is an independent and identically distributed (infinite) repetition of the type of state in question or, rather, in order for that phrase to make sense, an independent and identically distributed repetition of a certain ranking assessment of that type of state. Hence, a law is a certain kind of ranking function. This sounds weird, because a law thus turns into a kind of doxastic attitude. The literature on lawlikeness shows, however, that this is not so absurd a direction; if, besides explanatory power or support of counterfactuals, projectibility or inductive behavior are made defining features of laws, they are characterized by their epistemic role and thus get somehow entangled with our subjective states (see also Lange 2000, ch. 7, on the root commitment associated with laws). The main point, though, is that the ranking functions expressing deterministic laws are again of the objectifiable kind. So, there is a way of maintaining even within this account that laws obtain mind-independently.

In fact, according to what I have sketched, a deterministic law is the precise ranking analogue of a statistical law. De Finetti (1937) has proposed an ingenious way of eliminating objective probabilities and statistical laws by showing, in his famous representation theorem, that beliefs (i.e., subjective probabilities) about statistical laws (describing an infinite sequence of independent and identically distributed trials) are strictly equivalent to symmetric or exchangeable subjective probabilities for these trials and that experience makes these symmetric probabilities converge to the true statistical law. The eliminativist intention of the story is mostly dismissed today; rather, objective probabilities are taken seriously. Still, de Finetti's account has remained a paradigm story about the relation between subjective and objective probabilities.

I am mentioning all this because this paradigm story can be directly transferred to ranking theory. Let $\kappa$ be any ranking function for an infinite sequence of trials (= variables) which is regular and symmetric and according to which the outcome of a certain trial is not negatively relevant to the same outcome in the next trial. Then $\kappa$ is a unique mixture of deterministic laws for that sequence of trials in the above-mentioned sense, and experience makes $\kappa$ converge to the true deterministic law. (Cf. Spohn 2005a for all this, where I have treated only the simplest case of the infinite repetition of a binary variable or a trial having only two possible outcomes. With an additional condition, however, the results generalize to all variables taking finitely many values).

This may suffice as an overview over the basics of ranking theory and its elaboration into various directions; it got long enough. In a way, my overall argument in section "Further comparisons", pp. 335ff, when I shall make a bit more detailed comparative remarks about other members of the Baconian probability family, should be clear by now: Bayesian epistemology has enormous powers and virtues and rich details and ramifications. Small wonder that Pascal by far outstripped Bacon. In a nutshell, I have explained that many essential virtues can be duplicated in ranking theory; indeed, the duplications can stand on their own, having an independent significance. Bacon can catch up with Pascal. Of course, my rhetorical question will then be: Which other version of Baconian probability is able to come up with similar results?

Still, one might suspect that I can claim these successes only by turning Bacon into a fake Pascal. I have never left the Bayesian home, it may seem. Hence, one might even suspect that ranking theory is superfluous and may be reduced to the traditional Bayesian point of view. In other words, it is high time to study more closely the relation between probability and ranking theory. This will be our task in the next section.

## Ranks and Probabilities

The relation between probabilities and ranks is surprisingly complex and fascinating. I first turn to the more formal aspects of the comparison before discussing the philosophical aspects.

### *Formal Aspects*

The reader will have observed since long why ranks behave so much like probabilities. There is obviously a simple translation of probability into ranking theory: translate the sum of probabilities into the minimum of ranks, the product of probabilities into the sum of ranks, and the quotient of probabilities into the difference of ranks. Thereby, the probabilistic law of additivity turns into the law of disjunction, the probabilistic law of multiplication into the law of conjunction (for negative ranks), and the definition of conditional probabilities into the definition of conditional ranks. If the basic axioms and definitions are thus translated, then it is small wonder that the translation generalizes; take any probabilistic theorem, apply the above translation to it, and you are almost guaranteed to get a ranking theorem. This translation is obviously committed to negative ranks; therefore I always favored negative over positive ranks. However, the translation is not fool-proof; see, e.g., Spohn (1994a) for slight failures concerning conditional independence (between sets of variables) or Spohn (2005a) for slight differences concerning positive and non-negative instantial relevance. The issue is not completely cleared up.

Is there a deeper reason why this translation works so well? Yes, of course. The translation of products and quotients of probabilities suggests that negative ranks simply are the logarithm of probabilities (with respect to some base $< 1$). This does not seem to fit with the translation of sums of probabilities. But it does fit when the logarithmic base is taken to be some infinitesimal $i$ (since for two positive reals $x \leq y$ $i^x + i^y = i^{x\text{-}j}$ for some infinitesimal $j$). That is, we may understand ranks as real orders of magnitude of non-standard probabilities. This is the basic reason for the pervasive analogy.

Does this mean that ranking epistemology simply reduces to non-standard Bayesianism? This may be one way to view the matter. However, I do not particularly like this perspective. Bayesian epistemology in terms of non-standard reals is really non-standard. Even its great proponent, David Lewis, mentions the possibility only in passing (for the first time in 1980, p. 268). It is well known that both, non-standard analysis and its continuation as hyperfinite probability theory, have their intricacies of their own, and it is highly questionable from an epistemological point of view whether one should buy these intricacies. Moreover, even though this understanding of ranks is in principle feasible, it is nowhere worked out in detail. Such an elaboration should also explain the slight failures of the above translation. Hence, even formally the relation between ranks and non-standard probabilities is not fully clear. Finally, there are algebraic incoherencies. As long as the probabilistic law of additivity and the ranking law of disjunction are finitely restricted, there is no problem. However, it is very natural to conceive probability measures as σ-additive (although there is an argument about this point), whereas it is very natural to conceive of ranking functions as complete (as I have argued). This is a further disanalogy, which is not resolved by the suggested understanding of ranks.

All in all, I prefer to stick to the realm of standard reals. Ranking theory is a standard theory, and it should be compared to other standard theories. So, let us put the issue of hyperfinite probability theory to one side.

Let us instead pursue another line of thought. I have heavily emphasized that the fundamental point of ranking theory is to represent the statics and the dynamics of belief or of taking-to-be-true; it *is* the theory of belief. So, instead of inquiring the relation between ranks and probabilities we might as well ask the more familiar question about the relation between belief and probability.

This relation is well known to be problematic. One naive idea is that belief vaguely marks some threshold in probability, i.e., that *A* is believed iff its subjective probability is greater than $1 - \varepsilon$ for some small $\varepsilon$. But this will not do, as is hightlighted by the famous lottery paradox (see Kyburg 1961, p. 197, and Hempel (1962, pp. 163–166). According to this idea you may believe *A* and believe *B*, but fail to believe *A & B*. However, this amounts to saying that you do not know the truth table of conjunction, i.e., that you have not grasped conjunction at all. So, this idea is a bad one, as almost all commentators to the lottery paradox agree. One might think then about more complicated relations between belief and probability, but I confess not to have seen any convincing one.

The simplest escape from the lottery paradox is, of course, to equate belief with probability 1. This proposal faces two further problems, though. First, it seems intuitively inadequate to equate belief with maximal certainty in probabilistic terms; beliefs need not be absolutely certain. Secondly, but this is only a theoretical version of the intuitive objection, only belief expansion makes sense according to this proposal, but no genuine belief revision. Once you assign probability 1 to a proposition, you can never get rid of it according to all rules of probabilistic change. This is obviously inadequate; of course, we can give up previous beliefs and easily do so all the time.

Jeffrey's radical probabilism (1991) is a radical way out. According to Jeffrey, all subjective probabilities are regular, and his generalized conditionalization provides a dynamics moving within regular probabilities. However, Jeffrey's picture and the proposal of equating belief with probability 1 do not combine; then we would believe in nothing but the tautology. Jeffrey did not deny beliefs, but he indeed denied their relevance for epistemology; this is what the adjective 'radical' in effect signifies. He did not believe in any positive relation between belief and probability, and probability is all you need – a viable conclusion from the lottery paradox perhaps, though only as a last resort.

The point that probability theory cannot account for belief revision may apparently be dealt with by an expansion of the probabilistic point of view, namely by resorting to Popper measures. These take conditional probability as the basic notion, and thus probabilities conditional on propositions having absolute probability 0 may be well defined. That is, you may initially believe $A$, i.e., assign probability 1 to $A$, and still learn that $\overline{A}$, i.e., conditionalize w.r.t. $\overline{A}$, and thus move to posterior probabilities and even beliefs denying $A$. In this way, one can stick to the equation of belief with probability 1 and escape the above objection. Have we thus reached a stable position?

No, we have not. One point of Spohn (1986) was to rigorously show that AGM belief revision is just the qualitative counterpart of Popper measures. Conversely, this entails that the inability of AGM belief revision theory to model iterated belief revision, which I criticized in my (1988), holds for Popper measures as well. In fact, Harper (1976) was the first to note this problem vis à vis Popper measures, and thus I became aware of the problem and noticed the parallel.

Harper proposed quite a complicated solution to the problem that is, as far as I know, not well received; but it may be worth revisiting. My conclusion was a different one. If AGM belief revision theory is incomplete and has to be evolved into ranking theory, the probabilistic point of view needs likewise to get further expanded. We need something like probabilified ranks or ranked probabilities; it is only in terms of them that we can unrestrictedly explain iterated probabilistic change.

A ranking function associates with each rank a set of propositions having that rank. A ranked probability measure associates with each rank an ordinary probability measure. The precise definition is straightforward. Hence, I confined myself to mentioning the idea in my (1988, sect. 7); only in my (2005b) I took the trouble to explicitly introduce it. One should note, though, that as soon as one assumes the probability measures involved to be σ-additive, one again forces the ranks to be well-ordered (cf. Spohn 1986); this is why in my (2005b) only the probabilification of complete ranking functions is defined.

One may say that ranking theory thus ultimately reduces to probability theory. I find this misleading, however. What I have just sketched is rather a unification of probability and ranking theory; after all, we have employed genuine ranking ideas in order to complete the probabilistic point of view. The unification is indeed a powerful one; all the virtues of standard Bayesianism which I have shown to carry over to ranking theory hold for this unification as well. It provides a unified account of confirmation, of lawlikeness, even of causation. It appears to be a surprising, but most desirable wedding of Baconian and Pascalian probability. I shall continue on the topic in the next subsection.

The previous paragraphs again urge the issue of hyperfinite probability; ranked probabilities look even more like probabilities in terms of non-standard reals. However, I cannot say more than I already did; I recommend the issue for further investigation.[21] I should use the occasion for clarifying a possible confusion, though. McGee (1994, pp. 181ff.) showed that Popper measures correspond to non-standard probability measures in a specific way. Now, I have suggested that ranked probabilities do so as well. However, my (1986, 1988) together entail that ranked probabilities are more general than Popper measures. These three assertions do not fit together. Yet, the apparent conflict is easily dissolved. The correspondence proved by McGee is not a unique one. Different non-standard probability measures may correspond to the same Popper measure, just as different ranked probabilities may. Hence, if McGee says that the two approaches, Popper's and the non-standard one, "amount to the same thing" (p. 181), this is true only for the respects McGee is considering, i.e., w.r.t. conditional probabilities. It is not true for the wider perspective I am advocating here, i.e., w.r.t. probability dynamics.

## *Philosophical Aspects*

The relation between belief and probability is not only a formal issue, it is philosophically deeply puzzling. It would be disturbing if there should be two (or more) unrelated ways of characterizing our doxastic states. We must somehow come to grips with their relation.

The nicest option would be *reductionism*, i.e., reducing one notion to the other. This can only mean reducing belief to probability. As we have seen, however, this option seems barred by the lottery paradox. Another option is *eliminativism* as most ably defended in Jeffrey's radical probabilism also mentioned above. This option is certainly viable and most elegant. Still, I find it deeply unsatisfactory; it is unacceptable that our talk of belief should merely be an excusable error ultimately to be eliminated. Thus, both versions of *monism* seem excluded.

Hence, we have to turn to *dualism*, and then *interactionism* may seem the most sensible position. Of course, everything depends on the precise form of interaction between belief and probability. In Spohn (2005b) I had an argument with Isaac Levi whom I there described as the champion of interactionism. My general experience, though, is that belief and probability are like oil and water; they do not mix easily. Quite a different type of interactionism is represented by Hild (t.a.) who has many interesting things to say about how ranking and probability theory mesh, indeed how heavily ranking ideas are implicitly used in statistical methodology. I do not have space to assess this type of interactionism.

[21] For quite a different way of relating probabilities and ranks appealing neither to infinitesimals nor to Popperian conditional probabilities see Giang and Shenoy (1999).

When the fate of interactionism is unclear one might hope to return to reductionism and thus to monism, not in the form of reducing belief to probability, but in the form of *reducing both to something third*. This may be hyperfinite probability, or it may be ranked probabilities as suggested above. However, as already indicated, I consider this to be at best a formal possibility with admittedly great formal power of unification. Philosophically, I am not convinced. It is intuitively simply inadequate to equate belief with (almost) maximal probabilistic certainty, i.e., with probability 1 (minus an infinitesimal), even if this does not amount to unrevisability within these unifications. This intuition has systematic counterparts. For centuries, the behavioral connection of subjective probabilities to gambling and betting has been taken to be fundamental; many hold that this connection provides the only explanation of subjective probabilities. This fundamental connection does not survive these unifications. According to them, I would have to be prepared to bet my life on my beliefs; but this is true only of very few of my many beliefs. So, there are grave frictions that should not be plastered by formal means.

In view of all this, I have always preferred *separatism*, at least *methodologically*. If monism and interactionism are problematic, then belief and probability should be studied as two separate fields of interest. I sense the harshness of this position; this is why I am recommending it so far only as a methodological one and remain unsure about its ultimate status. However, the harshness is softened by the formal parallel which I have extensively exploited and which allows formal unification. Thus, separatism in effect amounts to *parallelism*, at least if belief is studied in ranking terms. Indeed, the effectiveness of the parallel sometimes strikes me as a pre-established harmony.

Thus, another moral to be drawn may perhaps be *structuralism*, i.e., the search for common structures. This is a strategy I find most clearly displayed in Halpern (2003). He starts with a very weak structure of degrees of belief that he calls plausibility measures and then discusses various conditions on those degrees that allow useful strengthenings of that structure such as a theory of conditioning, a theory of independence, a theory of expectation and integration, and so forth. Both, ranking and probability theory, but not only they are specializations of that structure and its various strengthenings. Without doubt, this is a most instructive procedure. Structuralism would moreover suggest that it is only those structures and not their specific realizations that matter. Halpern does not explicitly endorse this, and I think one should withstand it. For instance, one would thereby miss the essential purpose for which ranking theory was designed, namely the theory of belief. For this purpose, no less and no more than the ranking structure is required.

Hence, let me further pursue, in the spirit of methodological separatism, the philosophical comparison between ranks and standard probabilities. I have already emphasized the areas in which the formal parallel also makes substantial sense: inductive inference, confirmation, causation, etc. Let us now focus on three actual or apparent substantial dissimilarities, which in one or the other way concern the issue what our doxastic states have to do with reality.

The first aspect of this issue is the *truth connection*; ranks are related to truth in a way in which probabilities are not. This is the old point all over again.

Ranks represent beliefs that are true or false, whereas subjective probabilities do not represent beliefs and may be assessed in various ways, as well-informed, as reasonable, but never as true or false. Degrees of belief may perhaps conform to degrees of truthlikeness; however, it is not clear in the first place whether degrees of truthlikeness behave like probabilities (cf. Oddie 2001). Or degrees of belief may conform to what Joyce (1998) calls the norm of gradational accuracy from which he proceeds with an interesting argument to the effect that degrees of belief then have to behave like probabilities.[22] Such ideas are at best a weak substitute, however; they never yield an application of truth in probability theory as we have it in ranking theory.

This is a clear point in favor of ranking theory. And it is rich of consequences. It means that ranking theory, in contrast to probability theory, is able to connect up with traditional epistemology. For instance, Plantinga (1993, chs. 6 and 7) despairs of finding insights in Bayesianism he can use and dismisses it, too swiftly I find. This would have been different with ranking theory. The reason why ranking theory is connectible is obvious. Traditional epistemology is interested in knowledge, a category entirely foreign to probability theory; knowledge, roughly, is justified true belief and thus analyzed by notions within the domain of ranking theory. Moreover, the notion of justification has become particularly contested in traditional epistemology; one focal issue was then to give an account of the truth-conduciveness of reasons, again notions within the domain of ranking theory.

I am not claiming actual epistemological progress here. But I do claim an advantage of ranking over probability theory, I do claim that traditional epistemology finds in ranking theory adequate formal means for discussing its issues, and using such means is something I generally recommend as a formal philosopher.

The second aspect is the *behavioral connection*. Our doxastic states make some actions rational and others irrational, and our theories have to say which. Here, probability theory seems to have a clear advantage. The associated behavioral theory is, of course, decision theory with its fundamental principle of maximizing conditional expected utility. The power of this theory need not be emphasized here. Is there anything comparable on offer for ranking theory?

This appears excluded, for the formal reason that there is a theory of integration and thus of expectation in probabilistic, but none in ranking terms; this is at least what I had thought all along. However, the issue has developed. There are various remarkable attempts of stating a decision theory in terms of non-probabilistic or non-additive representations of degrees of belief employing the more general Choquet theory of integration.[23] Indeed, there is also one especially for ranking theory. Giang and Shenoy (2000) translate the axiomatic treatment of utility as it is

[22] Cf., however, Maher's (2002) criticism of Joyce's argument.

[23] Economists inquired the issue; see, e.g., Gilboa (1987), Schmeidler (1989), Jaffray (1989), Sarin and Wakker (1992) for early contributions, and Wakker (2005) for a recent one. The AI side concurs; see, e.g., Dubois and Prade (1995), Brafman and Tennenholtz (2000), and Giang and Shenoy (2005).

given by Luce and Raiffa (1957, sect. 2.5) in terms of simple and compound lotteries directly into the ranking framework, thus developing a notion of utility fitting to this framework. These attempts doubtlessly deserve further scrutiny (cf. also Halpern 2003, ch. 5).

Let me raise, though, another point relating to this behavioral aspect. Linguistic behavior is unique to humans and a very special kind of behavior. Still, one may hope to cover it by decision theoretic means, too. Grice's intentional semantics employs a rudimentary decision theoretic analysis, and Lewis' (1969) theory of conventions uses game (and thus decision) theoretic methods in a very sophisticated way. However, even Lewis' account of coordination equilibria may be reduced to a qualitative theory (in Lewis (1975) he explicitly uses only qualitative terminology). In fact, the most primitive linguistic behavioral law is the disquotation principle: if *a* seriously and sincerely utters "*p*", then *a* believes that *p*.[24] The point is that these linguistic behavioral laws and in particular the disquotation principle is stated in terms of belief. There is no probabilistic version of the disquotation principle, and it is unclear what it could be. The close relation between belief and meaning is obvious and undoubted, though perhaps not fully understood in the philosophy of language. I am not suggesting that there is a linguistic pragmatics in terms of ranking functions; there is hardly anything.[25] I only want to point out that the standing of ranking theory concerning this behavioral aspect is at least promising.

There is a third and final aspect, again apparently speaking in favor of probability theory. We do not only make decisions with the help of our subjective probabilities, we also do *statistics*. That is, we find a lot of *relative frequencies* in the world, and they are closely related to probabilities. We need not discuss here the exact nature of this relation. Concerning objective probabilities, it is extensively discussed in the debate about frequentism, and concerning subjective probabilities it is presumably best captured in Reichenbach's principle postulating that our subjective probabilities should rationally converge to the observed relative frequencies. What is clear, in any case, is that in some way or other relative frequencies provide a strong anchoring of probabilities in reality from which the powerful and pervasive application of statistical methods derives. Subjective probabilities are not simply free-floating in our minds.

For many years I thought that this is another important aspect in which ranking theory is inferior to probability theory. Recently, though, I have become more optimistic. Not that there would be any statistics in ranking terms[26]; I do not see ranks related to relative frequencies. However, a corresponding role is played by the notion of *exception* and thus by absolute frequencies. In section "Objective ranks?",

[24]If *a* speaks a foreign language, the principle takes a more complicated, but obvious form. There is also a disquotation principle for the hearer, which, however, requires a careful exchange of the hearer's and the speaker's role.

[25]See in particular Merin (2006, appendix B) and (2008) whose relevance-based pragmatics yields interesting results in probabilistic as well as in ranking-theoretic terms.

[26]However, I had already mentioned that Hild (t.a.) finds a much closer connection of probabilities and ranks within statistical methodology.

I left the precise account of objectifiable ranking functions in the dark. If one studies that account more closely, though, one finds that these objectifiable ranking functions, or indeed the laws as I have indicated them in section "Objective ranks?", are exception or fault counting functions. The rank assigned to some possible world by such a ranking function is just the number of exceptions from the law embodied in this function that occur in this world.

This is a dim remark so far, and here is not the place to elaborate on it. Still, I find the opposition of exceptions and relative frequencies appealing. Often, we take a type of phenomenon as more or less frequent, and then we apply our sophisticated statistical methodology to it. Equally often, we try to cover a type of phenomenon by a deterministic law, we find exceptions, we try to improve our law, we take recourse to a usually implicit ceteris paribus condition, etc. As far as I know, the methodology of the latter perspective is less sophisticated. Indeed, there is little theory. Mill's method of relevant variables, e.g., is certainly an old and famous attempt to such a theory (cf. its reconstruction in Cohen 1977, ch. 13). Still, both perspectives, the statistical and the deterministic one, are very familiar to us. What I am suggesting is that the deterministic perspective can be thoroughly described in terms of ranking theory.[27]

It would moreover be most interesting to attend to the vague borderline. Somewhere, we switch from one to the other perspective, from exceptions to small relative frequencies or the other way around. I am not aware of any study of this borderline, but I am sure it is worth getting inquired. It may have the potential of also illuminating the relation of belief and probability, the deterministic and the statistical attitude.

All these broad implications are involved in a comparison of ranks and probabilities. I would find it rather confusing to artificially combine them in some unified theory, be it hyperfinite or ranked probabilities. It is more illuminating to keep them separate. Also, I did not want to argue for any preference. I wanted to present the rich field of comparison in which both theories can show their great, though partially diverging virtues. There should be no doubt, however, that the driving force behind all these considerations is the formal *parallelism* which I have extensively used in section "The theory" (pp. 305ff) and explained in section "Formal aspects" (pp. 328ff).

## Further Comparisons

Let me close the paper with a number of brief comparative remarks about alternative accounts subsumable under the vague label 'Baconian probability'. I have already

[27] I attempted to substantiate this suggestion with my account of strict and ceteris paribus laws in Spohn (2002) and with my translation of de Finetti's representation theorem into ranking theory in Spohn (2005a). (New addendum: For the most recent ranking-theoretic account of ceteris paribus laws see Spohn (2014).)

made a lot of such remarks *en passant*, but it may be useful to have them collected. I shall distinguish between the earlier and usually more philosophical contributions on the one hand and the more recent, often more technical contributions from the computer science side on the other hand. The borderline is certainly fuzzy, and I certainly do not want to erect boundaries. Still, the centuries old tendency of specialization and of transferring problems from philosophy to special fields may be clearly observed here as well.

## *Earlier and Philosophical Literature*

It is perhaps appropriate to start with L. Jonathan Cohen, the inventor of the label. In particular his (1977) is an impressive document of dualism, indeed separatism concerning degrees of provability and degrees of probability or inductive (Baconian) and Pascalian probability. His work is, as far as I know, the first explicit and powerful articulation of the attitude I have taken here as well.[28]

However, his functions of inductive support are rather a preform of my ranking functions. His inductive supports correspond to my positive ranks. Cohen clearly endorsed the law of conjunction for positive ranks; see his (1970, pp. 21f. and p. 63). He also endorsed the law of negation; but he noticed its importance only in his (1977, pp. 177ff.), whereas in his (1970) it is well hidden as theorem 306 on p. 226. His presentation is a bit imperspicuous, though, since he is somehow attached to the idea that $\Box^i$, i.e., having an inductive support $\geq i$, behaves like iterable S4-necessity and since he even brings in first-order predicate calculus.

Moreover, Cohen is explicit on the relationality of inductive support; it is a two-place function relating evidence and hypothesis. Hence, one might expect to find a true account of conditionality. This, however, is not so. His conditionals behave like strict implication[29], a feature Lewis (1973, sect. 1.2–3) has already warned against. Moreover, Cohen discusses only laws of support with fixed evidence – with one exception, the consequence principle, as he calls it (1970, p. 62). Translated into my notation it says for a positive ranking function π that

$$\pi(C \mid A) \geq \pi(C \mid B) \text{ if } A \subseteq B, \tag{17.15}$$

which is clearly not a theorem of ranking theory. These remarks sufficiently indicate that the aspect so crucial for ranking functions is scarcely and wrongly developed in Cohen's work.

The first clear articulation of the basic Baconian structure is found, however, not in Cohen's work, but in Shackle (1949, 1969). His functions of potential surprise

[28]I must confess, though, that I had not yet noticed his work when I basically fixed my ideas on ranking functions in 1983.

[29]This is particularly obvious from Cohen (1970, p. 219, def. 5).

clearly correspond to my negative ranking functions; axiom (17.9) in (1969, p. 81) is the law of negation, and axiom (17.4) and/or (17.6) in (1969, p. 90) express the law of disjunction. At least informally, Shackle also recognizes the duality of positive and negative ranks. He is explicit that potential surprise expresses certainty of wrongness, i.e., disbelief, and that there is conversely certainty of rightness (1969, p. 74).

His general attitude, however, is not so decidedly dualistic as that of Cohen. His concern is rather a general account of uncertainty, and he insists that probability does not exhaust uncertainty. Probability is an appropriate uncertainty measure only if uncertainty is 'distributional', whereas potential surprise accounts for 'non-distributional' uncertainty. So, he also ends up with an antagonistic structure; but the intention was to develop two special cases of a general theory.

It is most interesting to see how hard Shackle struggles with an appropriate law of conjunction for negative ranks. The first version of his axiom 7 (1969, p. 80) claims, in our terminology, that

$$\kappa(A \cap B) = \max\{\kappa(A), \kappa(B)\}\,. \tag{17.16}$$

He accepts the criticism this axiom has met, and changes it into a second version (1969, p. 83), which I find must be translated into

$$\kappa(B) = \max\{\kappa(A), \kappa(B \mid A)\} \tag{17.17}$$

(and is hence no law of conjunction at all). He continues that it would be fallacious to infer that

$$\kappa(A \cap B) = \min[\max\{\kappa(A), \kappa(B \mid A)\}, \max\{\kappa(B), \kappa(A \mid B)\}]\,. \tag{17.18}$$

In (1969, ch. 24) he is remarkably modern in discussing "expectation of change of own expectations". I interpret his formula (i) on p. 199 as slightly deviating from the second version of his axiom 7 in claiming that

$$\kappa(A \cap B) = \max\{\kappa(A), \kappa(B \mid A)\}\,. \tag{17.19}$$

And on pp. 204f. he even considers, and rejects (for no convincing reason), the equation

$$\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A)\,, \tag{17.20}$$

i.e., our law of conjunction for negative ranks. In all these discussions, conditional degrees of potential surprise appear to be an unexplained primitive notion. So, Shackle may have been here on the verge of getting things right. On the whole, though, it seems fair to say that his struggle has not led to a clear result.

Isaac Levi has always pointed to this pioneering achievement of Shackle, and he has made his own use of it. In a way he did not develop Shackle's functions of potential surprise; he just stuck to the laws of negation and of disjunction for negative ranks. In particular, there is no hint of any notion of conditionalization. This is not to say that his epistemology is poorer than the one I have. Rather, he finds a place for Shackle's functions in his elaborated doxastic decision theory, more precisely, in his account of belief expansion. He adds a separate account of belief contraction, and with the help of what is called Levi's identity he can thus deal with every kind of belief change. He may even claim to come to grips with iterated change.[30] One may thus sense that his edifice is at cross-purposes with mine.

A fair comparison is hence a larger affair. I have tried to give it in Spohn (2005b). Let me only mention one divergence specifically related to ranking functions. Since Levi considers ranking functions as basically identical with Shackle's functions of potential surprise and since he sees the latter's role in expansion, he continuously brings ranking functions into the same restricted perspective. I find this inadequate. I rather see the very same structure at work at expansions as well as at contractions, namely the structure of ranks. Insofar I do not see any need of giving the two kinds of belief change an entirely different treatment.

This brings me to the next comparison, with AGM belief revision theory (cf. e.g., Gärdenfors 1988). I have already explained that I came to think of ranking theory as a direct response to the challenge of iterated belief revision for AGM belief revision theory, and I have explained how $A \rightarrow x$-conditionalization for ranks unifies and generalizes AGM expansion, revision, and contraction. One may wonder how that challenge was taken up within the AGM discussion. With a plethora of proposals (see Rott 2008), that partially ventilated ideas that I thought to have effectively criticized already in Spohn (1988) and that do not find agreement, as far as I see, with the exception of Darwiche and Pearl (1997). As mentioned, Hild and Spohn (2008) gives a complete axiomatization of iterated contraction. Whether it finds wider acceptance remains to be seen.

By no means, though, one should underestimate the richness of the AGM discussion, of which, e.g., Rott (2001) or Hanson (1999) give a good impression. A pertinent point is that ranking theory generalizes and thus simply sides with the standard postulates for revision and contraction (i.e., (K*1-8) and ($K^-$1-8) in Gärdenfors 1988, pp. 54–56 and 61–64). The ensuing discussion has shown that these postulates are not beyond criticism and that many alternatives are worth discussing (cf., e.g., Rott 2001, pp. 103ff., who lists three alternatives of K*7, nine of K*8, six of $K^-$7, and ten of $K^-$8). I confess I would not know how to modify ranking theory in order to do justice to such alternatives. Hence, a fuller comparison with AGM belief revision theory would have to advance a defense of the standard postulates against the criticisms related with the alternatives.

The point is, of course, relevant in the debate with Levi, too. He prefers what he calls mild contraction to standard AGM contraction that can be represented in

[30]Many aspects of his epistemology are already found in Levi (1967). The most recent statement is given in Levi (2004), where one also gets a good idea of the development of his thought.

ranking theory only as a form of iterated contraction. Again, one would have to discuss whether this representation is acceptable.

It is worth mentioning that the origins of AGM belief revision theory clearly lie in conditional logic. Gärdenfors' (1978) epistemic semantics for conditionals was a response to the somewhat unearthly similarity spheres semantics for counterfactuals in Lewis (1973), and via the so-called Ramsey test Gärdenfors' interest more and more shifted from belief in conditionals to conditional beliefs and thus to the dynamics of belief. Hence, one finds a great similarity in the formal structures of conditional logic and belief revision theory. In particular, Lewis' similarity spheres correspond to Gärdenfors' entrenchment relations (1988, ch. 4). In a nutshell, then, the progress of ranking theory over Lewis' counterfactual logic lies in proceeding from an ordering of counterfactuality (as represented by Lewis' nested similarity spheres) to a cardinal grading of disbelief (as embodied in negative ranking functions).[31]

Indeed, the origins reach back farther. Conditional logic also has a history, the earlier one being somewhat indeterminate. However, the idea of having an ordering of levels of counterfactuality or of far-fetchedness of hypotheses is explicitly found already in Rescher (1964). If $\pi$ is a positive ranking function taking only finitely many values $0, x_1, \ldots, x_m, \infty$, then $\pi^{-1}(\infty), \pi^{-1}(x_m), \ldots, \pi^{-1}(x_1), \pi^{-1}(0)$ is just a family of modal categories $M_0, \ldots, M_n$ $(n = m + 2)$, as Rescher (1964, pp. 47–50) describes it. His procedure on pp. 49 f. for generating modal categories makes them closed under conjunction; this is our law of conjunction for positive ranks. And he observes on p. 47 that all the negations of sentences in modal categories up to $M_{n\text{-}1}$ must be in $M_n = \pi^{-1}(0)$; this is our law of negation.

To resume, I cannot find an equivalent to the ranking account of conditionalization in all this literature. However, the philosophical fruits I have depicted in section "The theory", pp. 305ff., and also in section "Philosophical aspects", pp. 331ff., sprang from this account. Therefore, I am wondering to which extent this literature can offer similar fruits, and for all I know the answer tends to be negative.

## *More Recent Computer Science Literature*

In view of the exploding computer science literature on uncertainty since the 80s even the brief remarks in the previous subsection on the earlier times were disproportionate. However, it is important, I think, not to forget about the origins. My comparative remarks concerning the more recent literature must hence be even more cursory. This is no neglect, though, since Halpern (2003), in book length, provides comprehensive comparisons of the various approaches with an emphasis on those aspects (conditionalization, independence, etc.) that I take to be important, too. Some rather general remarks must do instead and may nevertheless be illuminating.

[31] For my ideas how to treat conditionals in ranking-theoretic terms see Spohn (2015).

In the computer science literature, ranking theory is usually subsumed under the heading "uncertainty" and "degrees of belief". This is not wrong. After all, ranks are degrees, and if (absolute) certainty is equated with unrevisability, revisable beliefs are uncertain beliefs. Still, the subsumption is also misleading. My concern was *not* to represent uncertainty and to ventilate alternative models of doing so. Thus stated, this would have been an enterprise with too little guidance. My concern was exclusively to statically and dynamically represent *ungraded* belief, and my observation was that this necessarily leads to the ranking structure. If this is so, then, as I have emphasized, all the philosophical benefits of having a successful representation of ungraded belief are conferred to ranking theory. By contrast, if one starts modeling degrees of uncertainty, it is always an issue (raised, for instance, by the lottery paradox vis à vis probability) to which extent such a model adequately captures belief and its dynamics. So, this is a principled feature that sets ranking theory apart from the entire uncertainty literature.

The revisability of beliefs was directly studied in computer science under headings like "default logic" or "nonmonotonic reasoning". This is another large and natural field of comparison for ranking theory. However, let me cut things short. The relation between belief revision theory and nonmonotonic reasoning is meticulously investigated by Rott (2001). He proved far-reaching equivalences between many variants on both sides. This is highly illuminating. At the same time, however, it is a general indication that the concerns that led me to develop AGM belief revision theory into ranking theory are not well addressed in these areas of AI. Of course, such lump-sum statements must be taken with caution.

The uncertainty literature has observed many times that the field of nonmonotonic reasoning is within its reach. Among many others, Pearl (1988, ch. 10) has investigated the point from the probabilistic side, and Halpern (2003, ch. 8) has summarized it from his more comprehensive perspective. This direction of inquiry is obviously feasible, but the reverse line of thought of deriving kinds of uncertainty degrees from kinds of nonmonotonic reasoning is less clear (though the results in Hild and Spohn (2008) about the measurement of ranks with via iterated contractions may be a step in the reverse direction).

So, let me return to accounts of uncertainty in a bit more detail, and let me take up *possibility theory* first. It originates from Zadeh (1978), i.e. from fuzzy set theory and hence from a theory of vagueness. Its elaboration in the book by Dubois and Prade (1988) and many further papers shows its wide applicability, but never denies its origin. So, it should at least be mentioned that philosophical accounts of vagueness (cf., e.g., Williamson 1994) have nothing much to do with fuzzy logic. If one abstracts from this interpretation, though, possibility theory is formally very similar to ranking theory. If *Poss* is a possibility measure, then the basic laws are:

$$Poss\,(\varnothing) = 0, Poss(W) = 1, \;\text{ and } Poss\,(A \cup B) = \max\,\{Poss(A), Poss(B)\}\,. \tag{17.21}$$

So far, the difference is merely one of scale. Full possibility 1 is negative rank 0, (im)possibility 0 is negative rank $\infty$, and translating the scales translates the characteristic axiom of possibility theory into the law of disjunction for negative ranks. Indeed, Dubois and Prade often describe their degrees of possibility in such a way that this translation fits not only formally, but also materially.

Hence, the key issue is again how conditionalization is treated within possibility theory. There is some uncertainty. First, there is the motive that also dominated Shackle's account of the functions of potential surprise, namely to keep possibility theory as an ordinal theory where degrees of possibility have no arithmetical meaning. Then the idea is to stipulate that

$$Poss\,(A \cap B) = \min\{Poss(A), Poss\,(B \mid A)\} = \min\{Poss(B), Poss\,(A \mid B)\}\,. \tag{17.22}$$

This is just Shackle's proposal (17.19). Hisdal (1978) proposed to go beyond (17.19) just by turning (17.22) into a definition of conditional possibility by additionally assuming that conditionally things should be as possible as possible, i.e., by defining $Poss(B \mid A)$ as the maximal degree of possibility that makes (17.22) true:

$$Poss\,(B \mid A) = \begin{Bmatrix} P\,(A \cap B)\,, \text{ if } Poss\,(A \cap B) < Poss(A) \\ 1, \text{ if } Poss\,(A \cap B) = Poss(A) \end{Bmatrix}. \tag{17.23}$$

Halpern (2003, Proposition 3.9.2, Theorem 4.4.5, and Corollary 4.5.8) entails that Bayesian net theory works also in terms of conditional possibility thus defined. Many things, though, do not work well. It is plausible that $Poss(B \mid A)$ is between the extremes 1 and $Poss(A \cap B)$. However, (17.23) implies that it can take only those extremes. This is unintelligible. (17.22) implies that, if neither $Poss(B \mid A)$ nor $Poss(A \mid B)$ is 1, they are equal, a strange symmetry. And so on. Such unacceptable consequences spread through the entire architecture.

However, there is a second way to introduce conditional possibilities (cf., e.g., Dubois and Prade 1998, p. 206), namely by taking numerical degrees of possibility seriously and defining

$$Poss\,(B \,||\, A) = Poss\,(A \cap B)\,/Poss(A). \tag{17.24}$$

This looks much better. Indeed, if we define $\kappa(A) = \log Poss(A)$, the logarithm taken w.r.t. some positive base $< 1$, then $\kappa$ is a negative ranking function such that also $\kappa(B \mid A) = \log Poss(B \,||\, A)$. Hence, (17.24) renders possibility and ranking theory isomorphic, and all the philosophical benefits may be gained in either terms. Still, there remain interpretational differences. If we are really up to degrees of belief and disbelief, then the ranking scale is certainly more natural; this is particularly clear when we look at the possibilistic analogue to two-sided ranking functions. My remarks about objectifiable ranking functions as fault counting functions would make no sense for a possibilistic scale. And so on. Finally, one must be aware that the philosophical benefits resulted from adequately representing *belief*. Hence, it is doubtful whether the formal structure suffices to maintain the benefits for alternative interpretations of possibility theory.

Let me turn to some remarks about (*Dempster-Shafer*) *DS belief functions*. Shafer (1976) built on Dempster's ideas for developing a general theory of evidence. He saw clearly that his theory covered all known conceptions of degrees of belief. This, and its computational manageability, explains its enormous impact. However, before entering any formal comparisons the first argument that should be settled is a philosophical one about the nature of evidence. There is the DS theory of evidence, and there is a large philosophical literature on observation and confirmation, Bayesianism being its dominant formal expression. I have explained why ranking theory and its account of reasons is a member of this family, too. Of course, this argument cannot even be started here. My impression, though, is that it is still insufficiently fought out, certainly hampered by disciplinary boundaries.

In any case, it is to be expected that DS belief functions and ranking functions are interpretationally at cross-purposes. This is particularly clear from the fact that negative ranking functions, like possibility measures or Shackle's functions of potential surprise, are formally a special case of DS belief functions; they are *consonant* belief functions as introduced in Shafer (1976, ch. 10). There, p. 219, Shafer says that consonant belief functions "are distinguished by their failure to betray even a hint of conflict in the evidence"; they "can be described as 'pointing in a single direction'." From the perspective of Shafer's theory of evidence this may be an adequate characterization. As a description of ranking functions, however, it does not make any sense whatsoever. This emphasizes that the intended interpretations diverge completely.

Even formally things do not fit together. We saw that the virtues of ranking theory depend on the specific behavior of conditional ranks. This does not generalize to DS belief functions. There is again an uncertainty how to conditionalize DS belief functions; there are two main variants (cf. Halpern 2003, p. 103 and p. 132, which I use as my reference book in the sequel). The central tool of Shafer's theory of evidence is the rule of combination proposed by Dempster (1967); it is supposed to drive the dynamics of DS belief functions. Combination with certain evidence is identical with one of the two variants of conditionalization (cf. Halpern 2003, p. 94). According to Shafer, other uncertain evidence is also to be processed by this rule. One might think, though, instead to handle it with Jeffrey's generalized conditionalization, which is indeed definable for both kinds of conditional belief functions (cf. Halpern 2003, p. 107). However, both kinds of Jeffrey conditionalization diverge from the rule of combination (cf. Halpern 2003, p. 107 and p. 114).

Indeed, this was my argument in Spohn (1990, p. 156) against formally equating ranking functions with consonant belief functions: Ranking dynamics is driven by a ranking analogue to Jeffrey conditionalization, but it cannot be copied by the rule of combination since the corresponding combinations move outside the realm of consonant belief functions. And, as I may add now, it does not help to let the dynamics of DS belief functions be driven by Jeffrey conditionalization instead of the rule of combination: Consonant belief functions are not closed under Jeffrey conditionalization as well, whereas ranking functions are thus closed.[32] I conclude that there is no formal subsumption of ranking functions under DS belief functions. Hence, their interpretations do not only actually diverge, they are bound to do so.

Smets' transferable belief model (cf., e.g., Smets 1998) proposes a still more general model for changing DS belief functions in terms of his so-called specializations. One should check whether it offers means for formally subsuming ranking functions under his model. Even if this would be possible, however, the interpretational concerns remain. Smets' specializations are so much wedded to Shafer's conception of evidence that any subsumption would appear artificial and accidental. The philosophical argument about the nature of evidence is even more pressing here.

A final remark: There is a bulk of literature treating doxastic uncertainty not in terms of a specific probability measure, but in terms of convex sets of probability measures. The basic idea behind this is that one's uncertainty is so deep that one is not even able to fix one's subjective probability. In this case, doxastic states may be described as sets of measures or in terms of probability intervals or in terms of lower and upper probabilities. Again, the multiple ways of elaborating this idea and their relations are well investigated (see again Halpern 2003). Indeed, DS belief functions, which provide a very general structure, emerges as generalizations of lower probabilities. Even they, though, do not necessarily transcend the probabilistic point of view, as Halpern (2003, p. 279) argues; DS belief functions are in a way tantamount to so-called inner measures. May we say, hence, that the alternative formal structures mentioned ultimately reduce to probabilism (liberalized in the way explained)? We may leave the issue open, though it is obvious that the liberal idea of uncertainty conceived as sets of subjective probabilities is, in substance, a further step away from the ideas determining ranking theory. Even if probabilism were successful in this way, as far as ranking theory is concerned we would only be thrown back to our comparative remarks in section "Ranks and probabilities", pp. 328ff.

We may therefore conclude that ranking theory is a strong independent pillar in that confusingly rich variety of theories found in the uncertainty literature. This conclusion is the only point of my sketchy comparative remarks. Of course, it is not to deny that the other theories serve other purposes well. It is obvious that we are still far from an all-purpose account of uncertainty or degrees of belief.

---

[32]Does this contradict the fact that ranking functions are equivalent to possibility measures (with their second kind of conditionalization), that possibility measures may be conceived as a special case of DS belief (or rather: plausibility) functions, and that Jeffrey conditionalization works for possibility measures as defined by Halpern (2003, p. 107)? No. The reason is that Jeffrey conditionalization for possibility measures is not a special case of Jeffrey conditionalization for DS belief functions in general. Cf. Halpern (2003, p. 107).

## References


Bacon, F. (1620), *Novum Organum*.

Brafman, R. I., & Tennenholtz, M. (2000). An axiomatic treatment of three qualitative decision criteria. *Journal of the Association of Computing Machinery, 47*, 452–482.

Carnap, R. (1950). *Logical foundations of probability*. Chicago: Chicago University Press.

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.

Darwiche, A., & Pearl, J. (1997). On the logic of iterated belief revision. *Artificial Intelligence, 89*, 1–29.

Dawid, A. P. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society B, 41*, 1–31.

de Finetti, B. (1937). La Prévision: Ses Lois Logiques, Ses Sources Subjectives. *Annales de l'Institut Henri Poincaré, 7*; engl. translation: (1964) Foresight: Its logical laws, its subjective sources. In: H. E. Kyburg, Jr., H. E., & Smokler (Eds.), *Studies in subjective probability* (pp. 93–158). New York: Wiley.

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

Dempster, A. P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society, Series B, 30*, 205–247.

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

Dubois, D., & Prade, H. (1995). *Possibility theory as basis for qualitative decision theory*. In: Proceedings of the 14th International Joint Conference on Artificial Iintelligence (IJCAI'95), Montreal, pp. 1925–1930.

Dubois, D., & Prade, H. (1998). Possibility theory: Qualitative and quantitative aspects. In D.M. Gabbay & P. Smets (Eds.), *Handbook of defeasible reasoning and uncertainty management systems* (Vol. 1) (pp. 169–226). Dordrecht: Kluwer.

Gabbay, D. M., et al. (Eds.). (1994). *Handbook of logic in artificial intelligence and logic programming, vol. 3, nonmonotonic reasoning and uncertainty reasoning*. Oxford: Oxford University Press.

Garber, D. (1980). Field and Jeffrey conditionalization. *Philosophy of Science, 47*, 142–145.

Gärdenfors, P. (1978). Conditionals and changes of belief. In I. Niiniluoto & R. Tuomela (Eds.), *The logic and epistemology of scientific change* (pp. 381–404). Amsterdam: North-Holland.

Gärdenfors, P. (1988). *Knowledge in flux*. Cambridge: MIT Press.

Geiger, D., & Pearl, J. (1990). On the logic of causal models. In R. D. Shachter, T. S. Levitt, J. Lemmer, & L. N. Kanal (Eds.), *Uncertainty in artificial intelligence 4* (pp. 3–14). Amsterdam: Elsevier.

Giang, P. G., & Shenoy, P. P. (1999). On transformations between probability and spohnian disbelief functions. In K. B. Laskey & H. Prade (Eds.), *Uncertainty in artificial intelligence* (Vol. 15, pp. 236–244). San Francisco: Morgan Kaufmann.

Giang, P. G., & 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.

Giang, P. G., & Shenoy, P. P. (2005). Two axiomatic approaches to decision making using possibility theory. *European Journal of Operational Research, 162*, 450–467.

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

Goldszmidt, M., & Pearl, J. (1996). Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence, 84*, 57–112.

Hacking, I. (1975). *The emergence of probability*. Cambridge: Cambridge University Press.

Halpern, J. Y. (2003). *Reasoning about uncertainty*. Cambridge: MIT Press.
Hansson, S.O. (ed.) (1997), *Special Issue on Non-Prioritized Belief Revision. Theoria* 63, 1–134.
Hansson, S. O. (1999). *A textbook of belief dynamics. Theory change and database updating*. Dordrecht: Kluwer.
Harper, W. L. (1976). Rational belief change, popper functions and counterfactuals. In W. L. Harper & C. A. Hooker (Eds.), *Foundations of probability theory, statistical inference, and statistical theories of science* (Vol. I, pp. 73–115). Dordrecht: Reidel.
Hempel, C. G. (1945). Studies in the logic of confirmation. *Mind, 54*, 1–26 + 97–121.
Hempel, C. G. (1962). Deductive-nomological vs. Statistical explanation. In H. Feigl & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science, vol. III, scientific explanation, space, and time* (pp. 98–169). Minneapolis: University of Minnesota Press.
Hild, M. (t.a.). *Introduction to induction: On the first principles of reasoning*, Manuscript.
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.
Hisdal, E. (1978). Conditional possibilities – independence and noninteractivity. *Fuzzy Sets and Systems, 1*, 283–297.
Hohenadel, S. (2013). *Efficient epistemic updates in rank-based belief networks*. Dissertation, University of Konstanz. See http://nbn-resolving.de/urn:nbn:de:bsz:352-250406
Huber, F. (2006). Ranking functions and rankings on languages. *Artificial Intelligence, 170*, 462–471.
Huber, F. (2007). The consistency argument for ranking functions. *Studia Logica, 86*, 299–329.
Hunter, D. (1991). Graphoids, semi-graphoids, and ordinal conditional functions. *International Journal of Approximate Reasoning, 5*, 489–504.
Jaffray, J.-Y. (1989). Linear utility theory for belief functions. *Operations Research Letters, 8*, 107–112.
Jeffrey, R. C. (1965). *The logic of decision.* Chicago: University of Chicago Press, 2nd ed. 1983.
Jeffrey, R. C. (1991). *Probability and the art of judgment*. Cambridge: Cambridge University Press.
Jensen, F. V. (2001). *Bayesian networks and decision graphs*. Berlin: Springer.
Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science, 65*, 575–603.
Joyce, J. (1999). *The foundations of causal decision theory*. Cambridge: Cambridge University Press.
Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. I). New York: Academic.
Krüger, L., et al. (1987). *The probabilistic revolution. Vol. 1: ideas in history, Vol. 2: ideas in the sciences*. Cambridge: MIT Press.
Kyburg, H. E., Jr. (1961). *Probability and the logic of rational belief*. Middletown: Wesleyan University Press.
Lange, M. (2000). *Natural laws in scientific practice*. Oxford: Oxford University Press.
Levi, I. (1967). *Gambling with truth*. New York: A. A. Knopf.
Levi, I. (2004). *Mild contraction: Evaluating loss of information due to loss of belief*. Oxford: Oxford University Press.
Lewis, D. (1969). *Convention: A philosophical study*. Cambridge: Harvard University Press.
Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.
Lewis, D. (1975). Languages and language. In K. Gunderson (Ed.), *Minnesota studies in the philosophy of science* (Vol. VII, pp. 3–35). Minneapolis: University of Minnesota Press.
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.
Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York: Wiley.
Maher, P. (2002). Joyce's argument for probabilism. *Philosophy of Science, 69*, 73–81.
McGee, V. (1994). Learning the impossible. In E. Eells & B. Skyrms (Eds.), *Probability and conditionals. Belief revision and rational decision* (pp. 179–199). Cambridge: Cambridge University Press.
Merin, A. (2006). *Decision theory of rhetoric*, book manuscript, to appear.

Merin, A. (2008). Relevance and reasons in probability and epistemic ranking theory. A study in cognitive economy. In: Forschungsberichte der DFG-Forschergruppe *Logik in der Philosophie* (Nr. 130). University of Konstanz.

Neapolitan, R. E. (1990). *Probabilistic reasoning in expert systems: Theory and algorithms*. New York: Wiley.

Oddie, G. (2001). Truthlikeness. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2001 Edition). http://plato.stanford.edu/archives/fall2001/entries/truthlikeness

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

Pearl, J. (2000). *Causality. Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Plantinga, A. (1993). *Warrant: The current debate*. Oxford: Oxford University Press.

Pollock, J. L. (1995). *Cognitive carpentry*. Cambridge: MIT Press.

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

Rescher, N. (1976). *Plausible reasoning*. Assen: Van Gorcum.

Rott, H. (2001). *Change, choice and inference: A study of belief revision and nonmonotonic reasoning*. Oxford: Oxford University Press.

Rott, H. (2008). Shifting priorities: Simple representations for twenty seven iterated theory change operators. In D. Makinson, J. Malinowski, & H. Wansing (Eds.), *Towards mathematical philosophy*. Dordrecht: Springer.

Sarin, R., & Wakker, P. P. (1992). A simple axiomatization of nonadditive expected utility. *Econometrica, 60*, 1255–1272.

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

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.

Shenoy, P. P. (1991). On Spohn's rule for revision of beliefs. *International Journal of Approximate Reasoning, 5*, 149–181.

Smets, P. (1998). The Transferable Belief Model for Quantified Belief Representation. In D.M. Gabbay & P. Smets (Eds.), *Handbook of defeasible reasoning and uncertainty management systems* (Vol. 1) (pp. 267–301). Dordrecht: Kluwer.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. Berlin: Springer, 2nd ed. 2000.

Spohn, W. (1976/78). *Grundlagen der Entscheidungstheorie*. Ph.D. thesis, University of Munich 1976, published: Kronberg/Ts.: Scriptor 1978, out of print, pdf-version at: http://www.uni-konstanz.de/FuF/Philo/Philosophie/philosophie/files/ge.buch.gesamt.pdf

Spohn, W. (1983). *Eine Theorie der Kausalität*, unpublished Habilitationsschrift, Universität München, pdf-version at: http://www.uni-konstanz.de/FuF/Philo/Philosophie/philosophie/files/habilitation.pdf

Spohn, W. (1986). The representation of Popper measures. *Topoi, 5*, 69–74.

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. (1991). A reason for explanation: Explanations provide stable reasons. In W. Spohn, B. C. van Fraassen, & B. Skyrms (Eds.), *Existence and explanation* (pp. 165–196). Dordrecht: Kluwer.

Spohn, W. (1993). Causal laws are objectifications of inductive schemes. In J. Dubucs (Ed.), *Philosophy of probability* (pp. 223–252). Dordrecht: Kluwer.

Spohn, W. (1994a). On the properties of conditional independence. In P. Humphreys (Ed.), *Patrick suppes: Scientific philosopher. Vol. 1: Probability and probabilistic causality* (pp. 173–194). Dordrecht: Kluwer.

Spohn, W. (1994b). On Reichenbach's principle of the common cause. In W. C. Salmon & G. Wolters (Eds.), *Logic, language, and the structure of scientific theories* (pp. 215–239). Pittsburgh: Pittsburgh University Press.

Spohn, W. (1999). Two coherence principles. *Erkenntnis, 50*, 155–175.

Spohn, W. (2001a). Vier Begründungsbegriffe. In T. Grundmann (Ed.), *Erkenntnistheorie. Positionen zwischen Tradition und Gegenwart* (pp. 33–52). Paderborn: Mentis.

Spohn, W. (2001b). Bayesian nets are all there is to causal dependence. In M. C. Galavotti, P. Suppes, & D. Costantini (Eds.), *Stochastic dependence and causality* (pp. 157–172). Stanford: CSLI Publications.

Spohn, W. (2002). Laws, ceteris paribus conditions, and the dynamics of belief. *Erkenntnis, 57*, 373–394; also in: Earman, J., Glymour, C., Mitchell, S. (Eds.). (2002). *Ceteris paribus laws* (pp. 97–118). Dordrecht: Kluwer.

Spohn, W. (2005a). Enumerative induction and lawlikeness. *Philosophy of Science, 72*, 164–187.

Spohn, W. (2005b). Isaac Levi's potentially surprising epistemological picture. In E. Olsson (Ed.), *Knowledge and inquiry: Essays on the pragmatism of Isaac Levi*. Cambridge: Cambridge University Press.

Spohn, W. (2006). Causation: An alternative. *British Journal for the Philosophy of Science, 57*, 93–119.

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

Spohn, W. (2014). The epistemic account of ceteris paribus conditions. *European Journal for the Philosophy of Science, 4*(2014), 385–408.

Spohn, W. (2015). Conditionals: A unified ranking-theoretic perspective. *Philosophers' Imprint 15*(1)1–30; see: http://quod.lib.umich.edu/p/phimp/3521354.0015.001/

Studeny, M. (1989). Multiinformation and the problem of characterization of conditional independence relations. *Problems of Control and Information Theory, 18*, 3–16.

Wakker, P. P. (2005). Decision-foundations for properties of nonadditive measures: General state spaces or general outcome spaces. *Games and Economic Behavior, 50*, 107–125.

Williamson, T. (1994). *Vagueness*. London: Routledge.

Zadeh, L. A. (1975). Fuzzy logics and approximate reasoning. *Synthese, 30*, 407–428.

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