Belief and Degrees of Belief

Franz Huber

1 Introduction

Degrees of belief are familiar to all of us. Our confidence in the truth of some propositions is higher than our confidence in the truth of other propositions. We are pretty confident that our computers will boot when we push their power button, but we are much more confident that the sun will rise tomorrow. Degrees of belief formally represent the strength with which we believe the truth of various propositions. The higher an agent's degree of belief for a particular proposition, the higher her confidence in the truth of that proposition. For instance, Sophia's degree of belief that it will be sunny in Vienna tomorrow might be .52, whereas her degree of belief that the train will leave on time might be .23. The precise meaning of these statements depends, of course, on the underlying theory of degrees of belief. These theories offer a formal tool to measure degrees of belief, to investigate the relations between various degrees of belief in different propositions, and to normatively evaluate degrees of belief.

The purpose of this book is to provide a comprehensive overview and assessment of the currently prevailing theories of degrees of belief. Degrees of belief are primarily studied in formal epistemology, but also in computer science and artificial intelligence, where they find applications in so-called expert systems and elsewhere. In the former case the aim is to adequately describe and, much more importantly, to normatively evaluate the epistemic state of an ideally rational agent. By employing the formal tools of logic and mathematics theories of degrees of belief allow a precise analysis that is hard to come by with traditional philosophical methods.

Different theories of degrees of belief postulate different ways in which degrees of beliefs are related to each other and, more generally, how epistemic states should be modeled. After getting a handle on the objects of belief in Section 2, we briefly survey the most important accounts in Section 3. Section 4 continues this survey by

F. Huber (✉)
Formal Epistemology Research Group, Zukunftskolleg and Department of Philosophy,
University of Konstanz, Germany
e-mail: franz.huber@uni-konstanz.de

focusing on the relation between belief and degrees of belief. Section 5 concludes this introduction by pointing at some relations to belief revision and nonmonotonic reasoning.

2 The Objects of Belief

Before we can investigate the relations between various degrees of belief, we have to get clear about the relata of the (degree of) belief relation. It is common to assume that belief is a relation between an epistemic agent at a particular time to an object of belief. Degree of belief is then a relation between a number, an epistemic agent at a particular time, and an object of belief. It is more difficult to state what the objects of belief are. Are they sentences or propositions expressed by sentences or possible worlds (whatever *these* are – see Stalnaker 2003) or something altogether different?

The received view is that the objects of belief are *propositions*, i.e. *sets of possible worlds* or truth conditions. A more refined view is that the possible worlds comprised by those propositions are *centered* at an individual at a given time (Lewis 1979). In that case the propositions are often called *properties*. Most epistemologists stay very general and assume only that there is a non-empty set of possibilities, W, such that exactly one element of W corresponds to the actual world. If the possibilities in W are centered, the assumption is that there is exactly one element of W that corresponds to your current time slice in the actual world (Lewis 1986 holds that this element not merely corresponds to, but *is* your current time slice in the actual world).

Centered propositions are needed to adequately represent self-locating beliefs such as Sophia's belief that she lives in Vienna, which may well be different from her belief that Sophia lives in Vienna (this is the case if Sophia does not believe that she is Sophia). Self-locating beliefs have important epistemological consequences (Elga 2000, Lewis, 2001), and centered propositions are ably argued by Egan (2006) to correspond to what philosophers have traditionally called *secondary qualities* (Locke 1690/1975). Lewis' (1979: 133ff) claim that the difference between centered and uncentered propositions plays little role in how belief and other attitudes are formally represented and postulated to behave in a rational way can only be upheld for synchronic constraints on the statics of belief. For diachronic constraints on the dynamics of belief this claim is false, because the actual centered world (your current time slice in the actual uncentered world) is continually changing as time goes by. We will bracket these complications, though, and assume that, unless noted otherwise, the difference between centered and uncentered possibilities and propositions has no effect on the topic at issue.

Propositions have a certain set-theoretic structure. The set of all possibibilities, W, is a proposition. Furthermore, if A and B are propositions, then so are the complement of A with respect to W, $W \setminus A = \overline{A}$, as well as the intersection of A and B, $A \cap B$. In other words, the set of propositions is a (finitary) *field* or *algebra* $\mathcal{A}$ over a non-empty set of possibilities W: a set that contains W and is closed under complementations and finite intersections. Sometimes the field of propositions, $\mathcal{A}$,

is not only assumed to be closed under finite, but under countable intersections. This means that $A_1 \cap \ldots \cap A_n \ldots$ is a proposition (an element of $\mathcal{A}$), if $A_1, \ldots, A_n \ldots$ are. Such a field $\mathcal{A}$ is called a σ-field. Finally, a field $\mathcal{A}$ is *complete* just in case the intersection $\bigcap \mathcal{B}$ of all sets in $\mathcal{B}$ is an element of $\mathcal{A}$, for each subset $\mathcal{B}$ of $\mathcal{A}$.

If Sophia believes (to some degree) that it will be sunny in Vienna tomorrow, but she does not believe (to the same degree) that it will not be not sunny in Vienna tomorrow, propositions cannot be the objects of Sophia's (degrees of) belief(s). After all, that it will be sunny in Vienna tomorrow and that it will not be not sunny in Vienna tomorrow is one and the same proposition. It is only expressed by two different, though logically equivalent sentences. For reasons like this some accounts take *sentences of a formal language* $\mathcal{L}$ to be the objects of belief. In that case the above mentioned set-theoretic structure translates into the following requirements: the tautological sentence τ is assumed to be in the language $\mathcal{L}$; and whenever α and β are in $\mathcal{L}$, then so is the negation of α, $\neg\alpha$, as well as the conjunction of α and β, $\alpha \wedge \beta$.

However, as long as logically equivalent sentences are required to be assigned the same degree of belief – and all accounts considered in this volume require this, because they are normative accounts – the difference between taking the objects of beliefs to be sentences of a formal language $\mathcal{L}$ or taking them to be propositions in a finitary field $\mathcal{A}$ is mainly cosmetic. Each formal language $\mathcal{L}$ induces a finitary field $\mathcal{A}$ over the set of all models or classical truth value assignments for $\mathcal{L}$, $Mod_{\mathcal{L}}$. It is simply the set of all propositions over $Mod_{\mathcal{L}}$ that are expressed by the sentences in $\mathcal{L}$. This set in turn induces a unique σ-field, viz. the smallest σ-field $\sigma(\mathcal{A})$ that contains $\mathcal{A}$ as a subset. It also induces a unique complete field, viz. the smallest complete field that contains $\mathcal{A}$ as a subset. In the present case where $\mathcal{A}$ is generated by $Mod_{\mathcal{L}}$, this complete field is the *powerset*, i.e. the set of all subsets, of $Mod_{\mathcal{L}}$, $\wp(Mod_{\mathcal{L}})$. Hence, if we start with a degree of belief function on a formal language $\mathcal{L}$, we automatically get a degree of belief function on the field $\mathcal{A}$ induced by $\mathcal{L}$. As we do not always get a language $\mathcal{L}$ from a field $\mathcal{A}$, the semantic framework of propositions is more general than the syntactic framework of sentences.

3 Theories of Degrees of Belief

We have started with the example of Sophia, whose degree of belief that it will be sunny in Vienna tomorrow equals .52. Usually degrees of belief are taken to be real numbers from the interval $[0, 1]$, but we will come across an alternative in Section 4. If the epistemic agent is certain that a proposition is true, her degree of belief for this proposition is 1. If the epistemic agent is certain that a proposition is false, her degree of belief for the proposition is 0. However, these are extreme cases. Usually we are neither certain that a proposition is true nor that it is false. That does not mean, though, that we are agnostic with respect to the question whether the proposition in question is true. Our belief that it is true may well be much stronger than that it is false. Degrees of belief quantify this strength of belief.

3.1 Subjective Probabilities

The best developed account of degrees of belief is the theory of subjective probabilities. On this view degrees of belief simply follow the laws of probability. Here is the standard definition due to Kolmogorov (1956). Let $\mathcal{A}$ be a field of propositions over the set of possibilities W. A function $\mathrm{Pr} : \mathcal{A} \to \mathfrak{R}$ from $\mathcal{A}$ into the set of real numbers, $\mathfrak{R}$, is a (*finitely additive* and *unconditional*) *probability* on $\mathcal{A}$ if and only if for all A and B in $\mathcal{A}$:

1. $\mathrm{Pr}(A) \geq 0$
2. $\mathrm{Pr}(W) = 1$
3. $\mathrm{Pr}(A \cup B) = \mathrm{Pr}(A) + \mathrm{Pr}(B)$ if $A \cap B = \emptyset$

The triple $\langle W, \mathcal{A}, \mathrm{Pr} \rangle$ is called a (*finitely additive*) *probability space*. If $\mathcal{A}$ is closed under countable intersections and thus a σ-field, and if Pr additionally satisfies

4. $\mathrm{Pr}(A_1 \cup \ldots \cup A_n \cup \ldots) = \mathrm{Pr}(A_1) + \cdots + \mathrm{Pr}(A_n) + \cdots$

Pr is a σ- or *countably additive* probability on $\mathcal{A}$ (Kolmogorov 1956: ch. 2 actually gives a different but equivalent definition – see e.g. Huber 2007a: sct. 4.1). In this case $\langle W, \mathcal{A}, \mathrm{Pr} \rangle$ is called a σ- or countably additive probability space.

A probability $\mathrm{Pr} : \mathcal{A} \to \mathfrak{R}$ on $\mathcal{A}$ is called *regular* just in case $\mathrm{Pr}(A) > 0$ for every non-empty A in $\mathcal{A}$. Let $\mathcal{A}^{\mathrm{Pr}}$ be the set of all propositions A in $\mathcal{A}$ with $\mathrm{Pr}(A) > 0$. The *conditional* probability $\mathrm{Pr}(\cdot \mid \circ) : \mathcal{A} \times \mathcal{A}^{\mathrm{Pr}} \to \mathfrak{R}$ on $\mathcal{A}$ (based on the unconditional probability $\mathrm{Pr} : \mathcal{A} \to \mathfrak{R}$ on $\mathcal{A}$) is defined for all A in $\mathcal{A}$ and all B in $\mathcal{A}^{\mathrm{Pr}}$ by the ratio

5. $\mathrm{Pr}(A \mid B) = \mathrm{Pr}(A \cap B) / \mathrm{Pr}(B)$

(Kolmogorov 1956, ch. 1, §4). The domain of the second argument place of $\mathrm{Pr}(\cdot \mid \circ)$ has to be restricted to $\mathcal{A}^{\mathrm{Pr}}$, since the fraction $\mathrm{Pr}(A \cap B) / \mathrm{Pr}(B)$ is not defined for $\mathrm{Pr}(B) = 0$. Note that $\mathrm{Pr}(\cdot \mid B) : \mathcal{A} \to \mathfrak{R}$ is a probability on $\mathcal{A}$, for every B in $\mathcal{A}^{\mathrm{Pr}}$. Other authors take conditional probability as primitive and define unconditional probability in terms of it (Hájek 2003).

What does it mean to say that Sophia's subjective probability for the proposition that tomorrow it will be sunny in Vienna equals .52? This is a difficult question. Let us first answer a different one. How do we measure Sophia's subjective probability for such a proposition? On one account Sophia's subjective probability for A is measured by her *betting ratio* for A, i.e. the highest price she is willing to pay for a bet that returns 1 Euro if A, and 0 otherwise. On a slightly different account Sophia's subjective probability for A is measured by her *fair* betting ratio for A, i.e. that number $r = b / (a + b)$ such that she considers the following bet to be fair: a Euros if A, and $-b$ Euros otherwise ($a, b \geq 0$ with inequality for at least one). As we may say it: Sophia considers it to be fair to bet you b to a Euros that A.

It is not irrational for Sophia to be willing to bet you 5.2 to 4.8 Euros that tomorrow it will be sunny in Vienna, but not be willing to bet you $520,000$ to $480,000$ Euros that this proposition is true. This uncovers one assumption of the measurement

in terms of (fair) betting ratios: the epistemic agent is assumed to be neither risk averse nor risk prone. Gamblers in the casino are risk prone: they pay more for playing roulette than the fair monetary value according to reasonable probabilities (this may be perfectly reasonable if the additional cost is what the gambler is willing to spend on the thrill she gets out of playing roulette). Sophia, on the other hand, is risk averse – and reasonably so! – when she refuses to bet you $100,000$ to $900,000$ Euros that it will be sunny in Vienna tomorrow, while she is happy to bet you 5 to 5 Euros that this proposition is true. After all, she might lose her standard of living along with this bet. Note that it does not help to say that Sophia's fair betting ratio for A is that number $r = b/(a + b)$ such that she considers the following bet to be fair: $1 - r = a/(a + b)$ Euros if A, and $-r = -b/(a + b)$ otherwise ($a, b \geq 0$ with inequality for at least one). Just as stakes of $1,000,000$ Euros may be too high for the measurement to work, stakes of 1 Euro may be too low.

Another assumption is that the agent's (fair) betting ratio for a proposition is independent of the truth values of the proposition. Obviously we cannot measure Sophia's subjective probability for the proposition that she will be happily married by the end of the week by offering her a bet that returns 1 Euro if she will, and 0 otherwise. Sophia's subjective probability for happily getting married by the end of the week will be fairly low (as a hardworking philosopher she does not have much time to date). However, assuming that happily getting married is something she highly desires, her betting ratio for this proposition will be fairly high.

Ramsey (1926) avoids the first assumption by using utilities instead of money. He avoids the second assumption by presupposing the existence of an "ethically neutral" proposition (a proposition whose truth or falsity does not affect the agent's utilities) which the agent takes to be just as likely to be true as she takes it to be false. For more see Hájek (2007).

Let us return to our question of what it means for Sophia to assign a certain subjective probability to a given proposition. It is one thing for Sophia to be willing to bet at particular odds or to consider particular odds as fair. It is another thing for Sophia to have a subjective probability of .52 that tomorrow it will be sunny in Vienna. Sophia's subjective probabilities are measured by, but not identical to her (fair) betting ratios. The latter are operationally defined and observable. The former are unobservable, theoretical entities that, following Eriksson and Hájek (2007), we should take as primitive.

The theory of subjective probabilities is not an adequate description of people's epistemic states (Kahneman et al. 1982). It is a normative theory that tells us how an ideally rational epistemic agent's degrees of belief should behave. So, why should such an agent's degrees of belief obey the probability calculus?

The *Dutch Book Argument* provides an answer to this question. (Cox's theorem, Cox 1946, and the representation theorem of measurement theory, Krantz et al. 1971, provide two further answers.) On its standard, pragmatic reading, the Dutch Book Argument starts with a link between degrees of belief and betting ratios as first premise. The second premise says that it is (pragmatically) defective to accept a series of bets which guarantees a sure loss. Such a series of bets is called a Dutch Book (hence the name "Dutch Book Argument"). The third premise is the

Dutch Book Theorem. Its standard, pragmatic version says that an agent's betting ratios obey the probability calculus if and only if an agent who has those betting ratios cannot be Dutch Booked (i.e. presented a series of bets each of which is acceptable according to those betting ratios, but whose combination guarantees a loss). From this it is inferred that it is (epistemically) defective to have degrees of belief that do not obey the probability calculus. Obviously this argument would be valid only if the link between degrees of belief and betting ratios were identity (in which case there were no difference between pragmatic and epistemic defectiveness) – and we have already seen that it is not.

Joyce (1998) attempts to vindicate probabilism by considering the *accuracy* of degrees of belief. The basic idea here is that a degree of belief function is (epistemically) defective if there exists an alternative degree of belief function that is more accurate in each possible world. The accuracy of a degree of belief $b(A)$ in a proposition A in a world w is identified with the distance between $b(A)$ and the truth value of A in w, where 1 represents truth and 0 represents falsehood. For instance, a degree of belief up to 1 in a true proposition is more accurate, the higher it is – and perfectly accurate if it equals 1. The overall accuracy of a degree of belief function b in a world w is then determined by the accuracy of the individual degrees of belief $b(A)$. Given some conditions on how to measure distance, Joyce is able to prove that a degree of belief function obeys the probability calculus if and only if there exists no alternative degree of belief function that is more accurate in each possible world (the only-if part is not explicitly mentioned in Joyce 1998, but needed for the argument to work and presented in Joyce's contribution to this volume). Therefore, degrees of belief should obey the probability calculus.

The objection to this attempt – due to Bronfman (manuscript) – that has attracted most attention starts by noting that Joyce's conditions on measures of inaccuracy do not determine a single measure, but a whole set of such measures. This would strengthen rather than weaken Joyce's argument, were it not for the fact that these measures differ in their recommendations as to which alternative degree of belief function a non-probabilistic degree of belief function should be replaced by. All of Joyce's measures of inaccuracy agree that an agent whose degree of belief function violates the probability axioms should adopt a probabilistic degree of belief function which is more accurate in each possible world. However, these measures may differ in their recommendation as to which particular probabilistic degree of belief function the agent should adopt. In fact, for each possible world, following the recommendation of one measure will leave the agent off less accurate according to some other measure. Why, then, should the epistemic agent move from her non-probabilistic degree of belief function to a probabilistic one in the first place?

In his contribution to this volume Joyce responds to this question and other objections. For more on Dutch Book Arguments, Joyce's non-pragmatic vindication of probabilism, and arguments for (non-) probabilism in general see Hájek's contribution to this volume.

We have discussed how to measure subjective probabilities, and why degrees of belief should obey the probability calculus. It is of particular epistemological interest how to update subjective probabilities when new information is received.

Whereas axioms 1–5 of the probability calculus are *synchronic* conditions on an ideally rational agent's degree of belief function, update rules are *diachronic* conditions that tell us how the ideally rational agent should revise her subjective probabilities when she receives new information of a certain format. If the new information comes in form of a certainty, probabilism is extended by

Update Rule 1 (Strict Conditionalization) *If* Pr : $\mathcal{A} \rightarrow \mathfrak{R}$ *is your subjective probability at time t, and between t and t' you learn E $\in \mathcal{A}$ and no logically stronger proposition, then your subjective probability at time t' should be* Pr$(\cdot \mid E) : \mathcal{A} \rightarrow \mathfrak{R}$.

Strict conditionalization thus says that the ideally rational agent's new subjective probability for a proposition *A* after becoming certain of *E* should equal her old subjective probability for *A* conditional on *E*.

Two questions arise. First, why should we update our subjective probabilities according to strict conditionalization? Second, how should we update our subjective probabilities when the new information is of a different format and we do not become certain of a proposition, but merely change our subjective probabilities for various propositions? Jeffrey (1983) answers the second question by what is now known as

Update Rule 2 (Jeffrey Conditionalization) *If* Pr : $\mathcal{A} \rightarrow \mathfrak{R}$ *is your subjective probability at time t, and between t and t' your subjective probabilities in the mutually exclusive and jointly exhaustive propositions* $E_1, \ldots, E_n, \ldots (E_i \in \mathcal{A})$ *change to* $p_1, \ldots, p_n, \ldots (p_i \in [0, 1])$ *with* $\sum_i p_i = 1$, *and the positive part of your subjective probability does not change on any superset thereof, then your subjective probability at time t' should be* Pr$'(\cdot) : \mathcal{A} \rightarrow \mathfrak{R}$, *where*

$$\text{Pr}'(\cdot) = \sum_i \text{Pr}(\cdot \mid E_i) \, p_i.$$

Jeffrey conditionalization thus says that the ideally rational agent's new subjective probability for *A* after changing her subjective probabilities for the elements E_i of a partition to p_i should equal the weighted sum of her old subjective probabilities for *A* conditional on the E_i, where the weights are the new subjective probabilities p_i for the elements of the partition.

One answer to the first question is the Lewis-Teller Dutch Book Argument for strict conditionalization that is analogous to the synchronic one discussed previously (Lewis 1999, Teller 1973). Its extension to Jeffrey conditionalization is presented in Armendt (1980) and discussed in Skyrms (1987). For more on the issue of diachronic coherence see Skyrms' contribution to this volume. As of now, there is no gradational accuracy argument for either strict or Jeffrey conditionalization. Other philosophers have provided arguments against strict (and, a fortiori, Jeffrey) conditionalization: van Fraassen (1989) holds that rationality does not require the adoption of a particular update rule (but see Kvanvig 1994), and Arntzenius (2003) uses, among others, the "shifting" nature of self-locating beliefs to argue against

strict conditionalization as well as against van Fraassen's reflection principle (van Fraassen 1995). The second feature used by Arntzenius (2003), called "spreading", is independent of self-locating beliefs. It will be mentioned again in Section 4.

In subjective probability theory complete ignorance of the epistemic agent with respect to a particular proposition A is often modeled by the agent's having a subjective probability of .5 for A as well as its complement $W \setminus A$. More generally, an agent with subjective probability Pr is said to be *ignorant* with respect to the partition $\{A_1, \ldots, A_n\}$ if and only if $\Pr(A_i) = 1/n$. The *principle of indifference* requires an agent to be ignorant with respect to a given partition (of "equally possible" propositions). It leads to contradictory results if the partition in question is not held fixed (see, for instance, the discussion of *Bertrand's paradox* in Kneale 1949). A more cautious version of this principle that is also applicable if the partition contains countably infinitely many elements is the *principle of maximum entropy*. It requires the agent to adopt one of those probability measures Pr as her degree of belief function over (the σ-field generated by) the countable partition $\{A_i\}$ that maximize the quantity

$$-\sum_i \Pr(A_i) \log \Pr(A_i).$$

The latter is known as the entropy of Pr with respect to the partition $\{A_i\}$. See Paris (1994).

Suppose Sophia has hardly any enological knowledge. Her subjective probability for the proposition that a Schilcher, an Austrian wine speciality, is a white wine might reasonably be .5, as might be her subjective probability that a Schilcher is a red wine. Contrast this with the following case. Sophia knows for sure that a particular coin is fair. That is, Sophia knows for sure that the objective chance of the coin landing heads as well as its objective chance of landing tails each equal .5. Under that assumption her subjective probability for the proposition that the coin will land heads on the next toss might reasonably be .5. Although Sophia's subjective probabilities are alike in these two scenarios, there is an important epistemological difference. In the first case a subjective probability of .5 represents complete ignorance. In the second case it represents substantial knowledge about the objective chances. (The principle that, roughly, one's initial subjective probabilities conditional on the objective chances should equal the objective chances is called the *principal principle* by Lewis 1980.)

Examples like these suggest that subjective probability theory does not provide an adequate account of degrees of belief, because it does not allow one to distinguish between ignorance and knowledge about chances. Interval-valued probabilities (Kyburg and Teng 2001, Levi 1980, van Fraassen 1990, Walley 1991) can be seen as a reply to this objection without giving up the probabilistic framework. In case the epistemic agent knows the objective chances she continues to assign sharp probabilities as usual. However, if the agent is ignorant with respect to a proposition A she will not assign it a subjective probability of .5 (or any other sharp value, for that matter). Rather, she will assign A a whole interval

$[a, b] \subseteq [0, 1]$ such that she considers any number in $[a, b]$ to be a legitimate subjective probability for A. The size $b - a$ of the interval $[a, b]$ reflects her ignorance with respect to A, that is, with respect to the partition $\{A, W \setminus A\}$. (As suggested by the last remark, if $[a, b]$ is the interval-probability for A, then $[1 - b, 1 - a]$ is the interval-probability for $W \setminus A$.) If Sophia were the enological ignoramus that we have previously imagined her to be, she would assign the interval $[0, 1]$ to the proposition that a Schilcher is a white wine. If she knows for sure that the coin she is about to toss has an objective chance of .5 of landing heads and she subscribes to the principal principle, $[.5, .5]$ will be the interval she assigns to the proposition that the coin, if tossed, will land heads.

When epistemologists say that knowledge implies belief (Steup 2006), they use a qualitative notion of belief that does not admit of degrees (except in the trivial sense that there is belief, disbelief, and suspension of judgment). The same is true for philosophers of language when they say that a normal speaker, on reflection, sincerely asserts to "A" only if she believes that A (Kripke 1979). This raises the question whether the qualitative notion of belief can be reduced to the quantitative notion of degree of belief. A simple thesis – known as the *Lockean thesis* – says that we should believe a proposition A just in case our degree of belief for A is sufficiently high ('should' takes wide scope over 'just in case'). Of course, the question is which threshold is sufficiently high. We do not want to require that we only believe those propositions whose truth we are certain of – especially if we follow Carnap (1962) and Jeffrey (2004) and require every reasonable subjective probability to be regular (otherwise we would not be allowed to believe anything except the tautology). We want to take into account our fallibilism, the fact that our beliefs often turn out to be false.

Given that degrees of belief are represented as subjective probabilities, this means that the threshold for belief should be less than 1. In terms of subjective probabilities, the Lockean thesis then says that an epistemic agent with subjective probability $\Pr : \mathcal{A} \to \mathfrak{R}$ should believe A in $\mathcal{A}$ just in case $\Pr(A) > 1 - \varepsilon$ for some $\varepsilon \in (0, 1]$. This, however, leads to the *lottery paradox* (Kyburg 1961, and, much clearer, Hempel 1962) as well as the *preface paradox* (Makinson 1965). For every threshold $\varepsilon \in (0, 1]$ there is a finite partition $\{A_1, \ldots, A_n\}$, $A_i \in \mathcal{A}$, and a reasonable subjective probability $\Pr : \mathcal{A} \to \mathfrak{R}$ such that $\Pr(A_i) > 1 - \varepsilon$ for all $i = 1, \ldots, n$, while $\Pr(A_1 \cap \ldots \cap A_n) < 1 - \varepsilon$.

For instance, let $\varepsilon = .02$ and consider a lottery with 100 tickets that is known for sure to be fair and such that exactly one ticket will win. Then it is reasonable, for every ticket $i = 1, \ldots, 100$, to assign a subjective probability of $1/100$ to the proposition that ticket i will win the lottery, T_i. We thus believe of each single ticket that it will lose, because $\Pr(W \setminus T_i) = .99 > 1 - .02$. Yet we also know for sure that exactly one ticket will win. So $\Pr(T_1 \cap \ldots \cap T_{100}) = 1 > 1 - .02$. We therefore believe both that at least one ticket will win, $T_1 \cap \ldots \cap T_{100}$, as well as of each individual ticket that it will not win: $W \setminus T_1, \ldots, W \setminus T_{100}$. Together these beliefs form a belief set that is inconsistent in the sense that its intersection is empty: $\bigcap \{T_1 \cap \ldots \cap T_{100}, W \setminus T_1, \ldots, W \setminus T_{100}\} = \emptyset$. Yet consistency (and deductive closure, which is implicit in taking propositions rather than sentences to be

the objects of belief) have been regarded as the minimal requirements on a belief set ever since Hintikka (1961).

The lottery paradox has led some people to reject the notion of belief altogether (Jeffrey 1970), whereas others have been led to the idea that belief sets need not be deductively closed (Foley 1992 and, especially, Foley's contribution to this volume). Still others have turned the analysis on its head and elicit a context-dependent threshold parameter ε from the agent's belief set. See Hawthorne and Bovens (1999) and, especially, Hawthorne's contribution to this volume.

Another view is to take the lottery paradox at face value and postulate two epistemic attitudes towards propositions – belief and degrees of belief – that are not reducible to each other. Frankish (2004) defends a particular version of this view. He distinguishes between a mind, where one unconsciously entertains beliefs, and a supermind, where one consciously entertains beliefs. For more see Frankish's contribution to this volume. Further discussion of the relation between belief and probabilistic degrees of belief can be found in Kaplan (1996) as well as Christensen (2004) and Maher (2006).

3.2 Dempster-Shafer Belief Functions

The theory of *Dempster-Shafer (DS) belief functions* (Dempster 1968, Shafer 1976) rejects the claim that degrees of belief can be measured by the epistemic agent's betting behavior. A particular version of the theory of DS belief functions is the *transferable belief model* (Smets and Kennes 1994). It distinguishes between two mental levels: the *credal* level and the *pignistic* level. Its twofold thesis is that fair betting ratios should indeed obey the probability calculus, but that degrees of belief, being different from fair betting ratios, need not. Degrees of belief need only satisfy the weaker DS principles. The idea is that whenever one is forced to bet on the pignistic level, degrees of belief are used to calculate fair betting ratios that satisfy the probability axioms (recall the Dutch Book Argument). These are then used to calculate the agent's expected utility for various acts (Savage 1972, Joyce 1999). However, on the credal level where one only entertains and quantifies various beliefs without using them for decision making, degrees of belief need not obey the probability calculus.

Whereas subjective probabilities are additive (axiom 3), DS belief functions $Bel : \mathcal{A} \to \mathfrak{R}$ are only *super-additive*, i.e. for all propositions A and B in $\mathcal{A}$:

6. $Bel(A) + Bel(B) \leq Bel(A \cup B)$ if $A \cap B = \emptyset$

In particular, the agent's the degree of belief for A and her degree of belief for $W \setminus A$ need not sum to 1.

What does it mean that Sophia's degree of belief for the proposition A is .52, if her degree of belief function is represented by a DS belief function $Bel : \mathcal{A} \to \mathfrak{R}$? According to one interpretation (Haenni and Lehmann 2003), the number $Bel(A)$ represents the strength with which A is supported by the epistemic agent's knowledge or belief base. It may well be that the agent's knowledge or belief base neither

supports A nor its complement $W \setminus A$, while it always maximally supports their disjunction, $A \cup \overline{A}$.

Recall the supposition that Sophia has hardly any enological knowledge. Under that assumption her knowledge or belief base will neither support the proposition that a Schilcher is a red wine, Red, nor will it support the proposition that a Schilcher is a white wine, $White$. However, Sophia may well be certain that a Schilcher is either a red wine or a white wine, $Red \cup White$. Hence her DS belief function Bel will be such that $Bel\,(Red) = Bel\,(White) = 0$ while $Bel\,(Red \cup White) = 1$.

On the other hand, Sophia knows for sure that the coin she is about to toss is fair. Hence her Bel will be such that $Bel\,(Heads) = Bel\,(Tails) = .5$. Thus we see that the theory of DS belief functions can distinguish between uncertainty and one form of ignorance. Indeed,

$$I\,(\mathrm{A}) = 1 - Bel\,(A_1) - \cdots - Bel\,(A_n) - \cdots$$

can be seen as a measure of the agent's ignorance with respect to the countable partition $\{A_1, \ldots, A_n, \ldots\}$ (the A_i may, for instance, be the values of a random variable such as *the price of a bottle of Schilcher in Vienna on November 21, 2007*).

Figuratively, a proposition A divides the agent's knowledge or belief base into three mutually exclusive and jointly exhaustive parts. A part that speaks in favor of A, a part that speaks against A (i.e. in favor of $W \setminus A$), and a part that neither speaks in favor of nor against A. $Bel\,(A)$ quantifies the part that supports A, $Bel\,(W \setminus A)$ quantifies the part that supports $W \setminus A$, and $I\,(A) = 1 - Bel\,(A) - Bel\,(W \setminus A)$ quantifies the part that neither supports A nor $W \setminus A$. Formally this is spelt out in terms of a (normalized) *mass function* on $\mathcal{A}$, a function $m : \mathcal{A} \to \mathfrak{R}$ such that for all propositions A in $\mathcal{A}$:

$$m\,(A) \geq 0$$
$$m\,(\emptyset) = 0$$
$$\sum_{B \in \mathcal{A}} m\,(B) = 1$$

A (normalized) mass function $m : \mathcal{A} \to \mathfrak{R}$ induces a DS belief function $Bel : \mathcal{A} \to \mathfrak{R}$ by defining, for each A in $\mathcal{A}$,

$$Bel\,(A) = \sum_{B \subseteq A} m\,(B).$$

The relation to subjective probabilities can now be stated as follows. Subjective probabilities require the epistemic agent to divide her knowledge or belief base into two mutually exclusive and jointly exhaustive parts: one that speaks in favor of A and one that speaks against A. That is, the neutral part has to be distributed among the positive and negative parts. Subjective probabilities can thus be seen as DS belief functions without ignorance.

A DS belief function $Bel : \mathcal{A} \to \mathfrak{R}$ induces a Dempster-Shafer *plausibility function* $P : \mathcal{A} \to \mathfrak{R}$, where for all A in $\mathcal{A}$,

$$P(A) = 1 - Bel(A).$$

Degrees of plausibility quantify that part of the agent's knowledge or belief base which is compatible with A, i.e. the part that supports A together with the part that neither supports A nor $W \setminus A$. In terms of the (normalized) mass function m inducing Bel this means that

$$P(A) = \sum_{B \cap A \neq \emptyset} m(B).$$

If and only if $Bel(A)$ and $Bel(W \setminus A)$ sum to less than 1, $P(A)$ and $P(W \setminus A)$ sum to more than 1. For more see Haenni's contribution to this volume.

The theory of DS belief functions is more general than the theory of subjective probabilities in the sense that the latter requires degrees of belief to be additive, while the former merely requires them to be super-additive. In another sense, though, the converse is true. The reason is that DS belief functions can be represented as *convex sets of probabilities* (Walley 1991). As not every convex set of probabilities can be represented as a DS belief function, sets of probabilities provide the most general framework we have come across so far.

An even more general framework is provided by Halpern's *plausibility measures* (Halpern 2003). These are functions $Pl : \mathcal{A} \to \mathfrak{R}$ such that for all propositions A and B in $\mathcal{A}$:

$Pl(\emptyset) = 0$
$Pl(W) = 1$

7. $Pl(A) \leq Pl(B)$ if $A \subseteq B$.

In fact, these are only the special cases of real-valued plausibility measures. While it is fairly uncontroversial that an agent's degree of belief function should obey Halpern's plausibility calculus, it is questionable whether his minimal principles are all there is to the rationality of degrees of belief. The resulting epistemology is, in any case, very thin.

3.3 Possibility Theory

Possibility theory (Dubois and Prade 1988) is based on fuzzy set theory (Zadeh 1978). According to the latter theory, an element need not belong to a set either completely or not at all, but may be a member of the set to a certain degree. For instance, Sophia may belong to the set of black haired women to degree .72, because her hair, although black, is sort of brown as well. This is represented by a *member-*

ship function $\mu_B : W \rightarrow [0, 1]$, where $\mu_B(w)$ is the degree to which woman $w \in W$ belongs to the set of black haired woman, B.

Furthermore, the degree $\mu_{\overline{B}}$ (Sophia) to which Sophia belongs to the set of women who do not have black hair, $\overline{B}$, equals $1 - \mu_B$ (Sophia). If $\mu_Y : W \rightarrow [0, 1]$ is the membership function for the set of young women, then the degree to which Sophia belongs to the set of black haired or young women, $B \cup Y$, is given by

$$\mu_{B \cup Y} \text{ (Sophia)} = \max \{\mu_B \text{ (Sophia)}, \mu_Y \text{ (Sophia)}\}.$$

Similarly, the degree to which Sophia belongs to the set of black haired young women, $B \cap Y$, is given by

$$\mu_{B \cap Y} \text{ (Sophia)} = \min \{\mu_B \text{ (Sophia)}, \mu_Y \text{ (Sophia)}\}.$$

μ_B (Sophia) is interpreted as the degree to which the vague statement "Sophia is a black haired woman" is true.

Degrees of truth belong to philosophy of language. They do not (yet) have anything to do with degrees of belief, which belong to epistemology. In particular, note that degrees of truth are usually considered to be *truth functional* (the truth value of a compound statement such as $A \wedge B$ is a function of the truth values of its constituent statements A and B; that is, the truth values of A and B determine the truth value of $A \wedge B$). This is the case for membership functions μ. Degrees of belief, on the other hand, are hardly ever considered to be truth functional. For instance, probabilities are not truth functional, because the probability of $A \cap B$ is not determined by the probability of A and the probability of B. That is, there is no function f such that for all probability spaces $\langle W, \mathcal{A}, \text{Pr} \rangle$ and all propositions A and B in $\mathcal{A}$:

$$\text{Pr}(A \cap B) = f(\text{Pr}(A), \text{Pr}(B))$$

Suppose I tell you that Sophia is tall. How tall is a tall woman? Is a woman with a height of 175 cm tall? Or does a woman have to be at least 178 cm in order to be tall? Although you know that Sophia is tall, your knowledge is incomplete due to the vagueness of the term "tall". Here possibility theory enters by equipping you with a (normalized) *possibility distribution*, a function $\pi : W \rightarrow [0, 1]$ with $\pi(\omega) = 1$ for at least one ω in W. The motivation for the latter requirement is that at least (in fact, exactly) one possibility is the actual possibility, and hence at least one possibility must be maximally possible. Such a possibility distribution $\pi : W \rightarrow [0, 1]$ on the set of possibilities W is extended to a *possibility measure* $\Pi : \mathcal{A} \rightarrow \mathfrak{R}$ on the field $\mathcal{A}$ over W by defining for each A in $\mathcal{A}$:

$\Pi(\emptyset) = 0$
$\Pi(A) = \sup \{\pi(\omega) : \omega \in A\}$

This entails that possibility measures $\Pi : \mathcal{A} \rightarrow \mathfrak{R}$ are *maxitive* (and hence *sub-additive*), i.e. for all A and B in $\mathcal{A}$,

8. $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$.

The idea is, roughly, that a proposition is at least as possible as all of the possibilities it comprises, and no more possible than the "most possible" possibility either. Sometimes, though, there is no most possible possibility (i.e. the supremum is no maximum). For instance, that is the case when the degrees of possibility are $1/2$, $3/4$, $7/8$, $\ldots$, $2^n - 1/2^n$, $\ldots$ In this case the degree of possibility for the proposition is the smallest number which is at least as great as all the degrees of possibilities of its elements. In our example this is 1. (As will be seen below, this is the main formal difference between possibility measures and unconditional ranking functions.)

We can define possibility measures without recourse to an underlying possibility distribution as functions $\Pi : \mathcal{A} \to \mathfrak{R}$ such that for all propositions A and B in $\mathcal{A}$:

$\Pi(\emptyset) = 0$
$\Pi(W) = 1$
$\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$

It is important to note, though, that the last clause is not well-defined for disjunctions or unions of infinitely many propositions (in this case one would have to use the supremum operation sup instead of the maximum operation max). The dual notion of a *necessity measure* $N : \mathcal{A} \to \mathfrak{R}$ is defined for all propositions A in $\mathcal{A}$ by

$$N(A) = 1 - \Pi\left(\overline{A}\right).$$

This implies that

$$N(A \cap B) = \min\{N(A), N(B)\}.$$

The latter equation can be used to start with necessity measures as primitive. Define them as functions $N : \mathcal{A} \to \mathfrak{R}$ such that for all propositions A and B in $\mathcal{A}$:

$N(\emptyset) = 0$
$N(W) = 1$
$N(A \cap B) = \min\{N(A), N(B)\}$

Then possibility measures $\Pi : \mathcal{A} \to \mathfrak{R}$ are obtained by the equation

$$\Pi(A) = 1 - N\left(\overline{A}\right).$$

Although the agent's epistemic state is completely specified by either Π or N, the agent's epistemic attitude towards a particular proposition A in $\mathcal{A}$ is only jointly specified by $\Pi(A)$ and $N(A)$. The reason is that, in contrast to probability theory, $\Pi(W \setminus A)$ is not determined by $\Pi(A)$. Thus, degrees of possibility (as well as degrees of necessity) are not truth functional either. The same is true for DS belief and DS plausibility functions as well as Halpern's plausibility measures.

In our example, let W_H be the set of values of the random variable $H = Sophia's$ *height in cm between 0 cm and 300 cm*, $W_H = \{0, \ldots, 300\}$. Let $\pi_H : W_H \rightarrow [0, 1]$ be your possibility distribution. It is supposed to represent your epistemic state concerning Sophia's body height, which contains your knowledge that she is tall. For instance, your π_H might be such that $\pi_H(n) = 1$ for any natural number $n \in [177, 185] \subset W$. In this case your degree of possibility for the proposition that Sophia is at least 177cm tall is

$$\Pi_H (H \geq 177) = \sup \{\pi_H(n) : n \geq 177\} = 1.$$

The connection to fuzzy set theory now is that your possibility distribution $\pi_H : W_H \rightarrow [0, 1]$, which is based on your knowledge that Sophia is tall, can be interpreted as the membership function $\mu_T : W_H \rightarrow [0, 1]$ of the set of tall woman. So the epistemological thesis of possibility theory is that your degree of possibility for the proposition that Sophia is 177 cm tall given the vague and hence incomplete knowledge that Sophia is tall equals the degree to which a 177 cm tall woman belongs to the set of tall woman. In more suggestive notation,

$$\pi_H (H = n \mid T) = \mu_T(n).$$

For more see the contribution to this volume by Dubois and Prade.

3.4 Summary

Let us summarize the accounts we have dealt with so far. Subjective probability theory requires degrees of belief to be additive. An ideally rational epistemic agent's subjective probability $\Pr : \mathcal{A} \rightarrow \Re$ is such that for any A and B in $\mathcal{A}$:

3. $\Pr(A) + \Pr(B) = \Pr(A \cup B)$ if $A \cap B = \emptyset$

The theory of DS belief functions requires degrees of belief to be super-additive. An ideally rational epistemic agent's DS belief function $Bel : \mathcal{A} \rightarrow \Re$ is such that for any A and B in $\mathcal{A}$:

6. $Bel(A) + Bel(B) \leq Bel(A \cup B)$ if $A \cap B = \emptyset$

Possibility theory requires degrees of belief to be maxitive and hence super-additive. An ideally rational epistemic agent's possibility measure $\Pi : \mathcal{A} \rightarrow \Re$ is such that for any A and B in $\mathcal{A}$:

7. $\Pi(A \cup B) = \max \{\Pi(A), \Pi(B)\}$

All of these functions are special cases of real-valued plausibility measures $Pl : \mathcal{A} \rightarrow \Re$, which are such that for all A and B in $\mathcal{A}$:

8. $Pl(A) \leq Pl(B)$ if $A \subseteq B$

We have seen that each of these accounts provides an adequate model for some epistemic situation (Halpern's plausibility measures do so trivially). We have further noticed that subjective probabilities do not give rise to a notion of belief that is consistent and deductively closed. Therefore the same is true for the more general DS belief functions and Halpern's plausibility measures. It has to be noted, though, that Roorda (1995) provides a definition of belief in terms of sets of probabilities. (As will be mentioned in the next section, there is notion of belief in possibility theory that is consistent and deductively closed in a finite sense.)

Moreover, we have seen arguments for the thesis that degrees of belief should obey the probability calculus. Smets (2002) tries to justify the corresponding thesis for DS belief functions. To the best of my knowledge nobody has yet published an argument for the thesis that degrees of belief should obey Halpern's plausibility calculus (not just in the sense that only plausibility measures are reasonable degree of belief functions, but in the sense that *all* and only plausibility measures are reasonable degree of belief functions.) I am not aware of an argument for the corresponding thesis for possibility measures either. However, there exists such an argument for the formally similar ranking functions. These functions also give rise to a notion of belief that is consistent and deductively closed. They are the topic of. the next section.

4 Belief, Degrees of Belief, and Ranking Functions

Subjective probability theory as well as the theory of DS belief functions take the objects of belief to be propositions. Possibility theory does so only indirectly, though possibility measures on a field of propositions A can also be defined without recourse to a possibility distribution on the underlying set of possibilities W. A possibility ω in W is a complete and consistent description of what the world may look like relative to the expressive power of W. W may contain two possibilities: according to ω_1 it will be sunny in Vienna tomorrow, according to ω_2 it will not. On the other end of the spectrum, W may comprise grand possible worlds à la Lewis (1986).

We usually do not know which of the possibilities in W corresponds to the actual world. Otherwise these possibilities would not be genuine possibilities for us, and our degree of belief function would collapse into the truth value assignment corresponding to the actual world. All we usually know for sure is that there is exactly one possibility which corresponds to the actual world. However, to say that we do not know which possibility that is does not mean that all possibilities are on a par. Some of them will seem really far-fetched, while others will strike us as more reasonable candidates for the actual possibility.

This gives rise to the following consideration. We can *partition* the set of possibilities, that is, form sets of possibilities that are mutually exclusive and jointly exhaustive. Then we can *order* the cells of this partition according to their plausibility. The first cell in this ordering contains the possibilities that we take to be the most reasonable candidates for the actual possibility. The second cell contains the possibilities which we take to be the second most reasonable candidates. And so on.

If you are still equipped with your possibility distribution from the preceding section you can use your degrees of possibility for the various possibilities to obtain such an *ordered partition*. Note, though, that an ordered partition – in contrast to your possibility distribution – contains no more than ordinal information. While your possibility distribution enables you to say *how* possible you take a possibility to be, an ordered partition only allows you to say that one possibility ω_1 is more plausible than another ω_2. In fact, an ordered partition does not even let you say that the difference between your plausibility for w_1 (say, tomorrow the temperature in Vienna will be between 15°C and 20°C) and for w_2 (say, tomorrow the temperature in Vienna will be between 20°C and 25°C) is smaller than the difference between your plausibility for w_2 and for the far-fetched w_3 (say, tomorrow the temperature in Vienna will be between 45°C and 50°C).

This takes us directly to *ranking theory* (Spohn 1988; 1990), which goes one step further. Rather than merely ordering the possibilities in W, a *pointwise ranking function* $\kappa : W \rightarrow N \cup \{\infty\}$ additionally assigns natural numbers to the (cells of) possibilities. These numbers represent the degree to which an ideally rational epistemic agent disbelieves the various possibilities in W. The result is a *numbered partition* of W,

$$\kappa^{-1}(0), \kappa^{-1}(1), \ldots, \kappa^{-1}(n) = \{\omega \in W : \kappa(\omega) = n\}, \ldots, \kappa^{-1}(\infty).$$

The first cell $\kappa^{-1}(0)$ contains the possibilities the agent does not disbelieve (which does not mean that she believes them). The second cell $\kappa^{-1}(1)$ is the set of possibilities the agent disbelieves to degree 1. And so on. It is important to note that, except for $\kappa^{-1}(0)$, the cells $\kappa^{-1}(n)$ may be empty, and so would not appear at all in the corresponding ordered partition. $\kappa^{-1}(0)$ must not be empty, though. The reason is that one cannot consistently disbelieve everything.

More precisely, a function $\kappa : W \rightarrow N \cup \{\infty\}$ from a set of possibilities W into the set of natural numbers extended by ∞, $N \cup \{\infty\}$, is a (normalized) pointwise ranking function just in case $\kappa(\omega) = 0$ for at least one ω in W, i.e. just in case $\kappa^{-1}(0) \neq \emptyset$. The latter requirement says that the agent should not disbelieve every possibility. It is justified, because she knows for sure that one possibility is the actual one. A pointwise ranking function $\kappa : W \rightarrow N \cup \{\infty\}$ on W induces a *ranking function* $\varrho : \mathcal{A} \rightarrow N \cup \{\infty\}$ on a field of propositions $\mathcal{A}$ over W by defining for each A in $\mathcal{A}$,

$$\varrho(A) = \min\{\kappa(\omega) : \omega \in A\} \quad (= \infty \text{ if } A = \emptyset).$$

This entails that ranking functions $\varrho : \mathcal{A} \rightarrow N \cup \{\infty\}$ are *(finitely) minimitive* (and hence sub-additive), i.e. for all propositions A and B in $\mathcal{A}$,

9. $\varrho(A \cup B) = \min\{\varrho(A), \varrho(B)\}.$

As in the case of possibility theory, (finitely minimitive and unconditional) ranking functions can be directly defined on a field of propositions $\mathcal{A}$ over a set of possibilities W as functions $\varrho : \mathcal{A} \rightarrow N \cup \{\infty\}$ such that for all A and B in $\mathcal{A}$:

$$\varrho\,(\emptyset) = \infty$$
$$\varrho\,(W) = 0$$
$$\varrho\,(A \cup B) = \min\{\varrho\,(A)\,,\varrho\,(B)\}$$

The triple $\langle W, \mathcal{A}, \varrho \rangle$ is a (finitely minimitive) *ranking space*. Suppose $\mathcal{A}$ is closed under countable/complete intersections (and thus a σ-/complete field). Suppose further that ϱ additionally satisfies, for every countable/possibly uncountable $\mathcal{B} \subseteq \mathcal{A}$,

$$\varrho\,(\mathcal{B}) = \min\{\varrho\,(A) : A \in \mathcal{B}\}\,.$$

Then ϱ is a *countably/completely minimitive* ranking function, and $\langle W, \mathcal{A}, \varrho \rangle$ is a countably/completely minimitive ranking space. Finally, a ranking function ϱ on $\mathcal{A}$ is *regular* just in case $\varrho\,(A) < \infty$ for every non-empty or consistent proposition A in $\mathcal{A}$. For more see Huber (2006), which discusses under which conditions ranking functions on fields of propositions induce pointwise ranking functions on the underlying set of possibilities.

Let us pause for a moment. The previous paragraphs introduce a lot of terminology for something that seems to add only little to what we have already discussed. Let the necessity measures of possibility theory assign natural instead of real numbers in the unit interval to the various propositions so that ∞ instead of 1 represents maximal necessity and maximal possibility. Then the axioms for necessity measures become:

$$N\,(\emptyset) = 0, \quad N\,(W) = \infty, \quad N\,(A \cap B) = \min\{N\,(A)\,,N\,(B)\}$$

Now think of the rank of a proposition A as the degree of necessity of its negation $W \setminus A$, $\varrho\,(A) = N\,(W \setminus A)$. Seen this way, finitely minimitive ranking functions are a mere terminological variation of necessity measures:

$$\varrho\,(\emptyset) = N\,(W) = \infty$$
$$\varrho\,(W) = N\,(\emptyset) = 0$$
$$\varrho\,(A \cup B) = N\left(\overline{A} \cap \overline{B}\right) = \min\left\{N\left(\overline{A}\right), N\left(\overline{B}\right)\right\} = \min\{\varrho\,(A)\,,\varrho\,(B)\}$$

(If we take necessity measures as primitive rather than letting them be induced by possibility measures, and if we continue to follow the rank-theoretic policy of adopting a well-ordered range, we can obviously also define countably and completely minimitive necessity measures.) Of course, the fact that (finitely minimitive and unconditional) ranking functions and necessity measures are formally alike does not mean that their interpretations are the same. The latter is the case, though, when we compare ranking functions and Shackle's degrees of *potential surprise* (Shackle 1949; 1969). (These degrees of potential surprise have made their way into philosophy mainly through the work of Isaac Levi – see Levi 1967a; 1978.) So what justifies devoting a whole section to ranking functions?

Shackle's theory lacks a notion of *conditional* potential surprise. Shackle (1969: 79ff) seems to assume a notion of conditional potential surprise as primitive that appears in his axiom 7. This axiom further relies on a connective that behaves like conjunction except that it is not commutative and is best interpreted as "A followed by B". Axiom 7, in its stronger version from p. 83, seems to say that the degree of potential surprise of "A followed by B" is the greater of the degree of potential surprise of A and the degree of potential surprise of B given A,

$$\varsigma\,(A \text{ followed by } B) = \max\,\{\varsigma\,(A), \varsigma\,(B \mid A)\},$$

where ς is the measure of potential surprise. Spohn's contribution to this volume also discusses Shackle's struggle with the notion of conditional potential surprise.

Possibility theory, on the other hand, offers two notions of conditional possibility (Dubois and Prade 1988). The first notion of conditional possibility is obtained by the equation

$$\Pi\,(A \cap B) = \min\,\{\Pi\,(A), \Pi\,(B \mid A)\}.$$

It is mainly motivated by the desire to have a notion of conditional possibility that makes also sense if possibility does not admit of degrees, but is a merely comparative notion. The second notion of conditional possibility is obtained by the equation

$$\Pi\,(A \cap B) = \Pi\,(A)\,\Pi\,(B \parallel A).$$

The inspiration for this notion seems to come from probability theory. While none of these two notions is the one we have in ranking theory, Spohn's contribution to this volume (relying on Halpern 2003) shows that, by adopting the second notion of conditional possibility, one can render possibility theory isomorphic to real-valued ranking functions. For reasons explained below, I prefer to stick to ranking functions taking only natural numbers as values, though – and for the latter there is just one good notion of conditional ranks.

The *conditional* ranking function $\varrho\,(\cdot \mid \cdot) : \mathcal{A} \times \mathcal{A} \rightarrow N \cup \{\infty\}$ (based on the unconditional ranking function $\varrho : \mathcal{A} \rightarrow N \cup \{\infty\}$) is defined for all A and B in $\mathcal{A}$ with $A \neq \emptyset$ as

$$\varrho\,(A \mid B) = \varrho\,(A \cap B) - \varrho\,(B),$$

where $\infty - \infty =\cdot 0$. Further stipulating $\varrho\,(\emptyset \mid B) = \infty$ for all B in $\mathcal{A}$ guarantees that $\varrho\,(\cdot \mid B) : \mathcal{A} \rightarrow N \cup \{\infty\}$ is a ranking function, for every B in $\mathcal{A}$. It is, of course, also possible to take conditional ranking functions as primitive and to define (unconditional) ranking functions in terms of them.

The number $\varrho\,(A)$ represents the agent's degree of disbelief for the proposition A. If $\varrho\,(A) > 0$, the agent disbelieves A to a positive degree. Therefore, on pain of inconsistency, she cannot also disbelieve $W \setminus A$ to a positive degree. In other words,

for every proposition A in $\mathcal{A}$, at least one of A and $W \setminus A$ has to be assigned rank 0. If $\varrho(A) = 0$, the agent does not disbelieve A to any positive degree. This does not mean, however, that she believes A to a positive degree – the agent may suspend judgement and assign rank 0 to both A and $W \setminus A$. Belief in a proposition is thus characterized as disbelief in its negation.

For each ranking function $\varrho : \mathcal{A} \to N \cup \{\infty\}$ we can define a corresponding *belief function* $\beta_\varrho : \mathcal{A} \to Z \cup \{\infty\} \cup \{-\infty\}$ that assigns positive numbers to those propositions that are believed, negative numbers to those that are disbelieved, and 0 to those with respect to which the agent suspends judgement:

$$\beta_\varrho(A) = \varrho(W \setminus A) - \varrho(A)$$

Each ranking function $\varrho : \mathcal{A} \to N \cup \{\infty\}$ induces a *belief set*:

$$\mathcal{B}_\varrho = \left\{ A \in \mathcal{A} : \varrho\left(\overline{A}\right) > 0 \right\} = \left\{ A \in \mathcal{A} : \varrho\left(\overline{A}\right) > \varrho(A) \right\} = \left\{ A \in \mathcal{A} : \beta_\varrho(A) > 0 \right\}$$

$\mathcal{B}$ is the set of all propositions the agent believes to some positive degree or, equivalently, whose complements she disbelieves to a positive degree. The belief set $\mathcal{B}_\varrho$ induced by a ranking function ϱ is consistent and deductively closed (in the finite sense). The same is true for the belief set induced by a possibility measure $\Pi : \mathcal{A} \to \mathfrak{R}$,

$$\mathcal{B}_\Pi = \left\{ A \in \mathcal{A} : \Pi\left(\overline{A}\right) < 1 \right\} = \{ A \in \mathcal{A} : N_\Pi(A) > 0 \}.$$

If ϱ is a countably/completely minimitive ranking function, then the belief set $\mathcal{B}_\varrho$ induced by ϱ is consistent and deductively closed in the following countable/complete sense: $\bigcap \mathcal{C} \neq \emptyset$ for every countable/possibly uncountable $\mathcal{C} \subseteq \mathcal{B}_\varrho$; and $A \in \mathcal{B}_\varrho$ whenever $\bigcap \mathcal{C} \subseteq A$ for any countable/possibly uncountable $\mathcal{C} \subseteq \mathcal{B}_\varrho$ and any $A \in \mathcal{A}$. Ranking theory thus offers a link between belief and degrees of belief that is preserved when we move from the finite to the countably or uncountably infinite case. As shown by the example in Section 3.3, this is not the case for possibility theory. (Of course, as indicated above, the possibility theorist can copy ranking theory by taking necessity measures as primitive and by adopting a well-ordered range).

As for subjective probabilities there are rules for updating one's epistemic state represented by a ranking function. In case the new information comes in form of a certainty, ranking theory's counterpart to probability theory's strict conditionalization is

Update Rule 3 (Plain Conditionalization) *If* $\varrho : \mathcal{A} \to N \cup \{\infty\}$ *is your ranking function at time t, and between t and t' your learn* $E \in \mathcal{A}$ *and no logically stronger proposition, then your ranking function at time t' should be* $\varrho(\cdot \mid E) : \mathcal{A} \to N \cup \{\infty\}$.

If the new information merely changes your ranks for various propositions, ranking theory's counterpart to probability theory's Jeffrey conditionalization is

Update Rule 4 (Spohn Conditionalization) *If $\varrho : \mathcal{A} \to N \cup \{\infty\}$ is your ranking function at time t, and between t and t' your ranks in the mutually exclusive and jointly exhaustive propositions $E_1, \ldots, E_m, \ldots (E_i \in \mathcal{A})$ change to $n_1, \ldots, n_m, \ldots$ ($n_i \in N \cup \{\infty\}$) with $\min_i \{n_i\} = 0$, and the finite part of your ranking function does not change on any superset thereof, then your ranking function at time t' should be $\varrho' : \mathcal{A} \to N \cup \{\infty\}$, where*

$$\varrho'(\cdot) = \min_i \{\varrho(\cdot \mid E_i) + n_i\}.$$

As the reader will have noticed by now, whenever we substitute 0 for 1, ∞ for 0, min for $+$, $+$ for $\times$, and $>$ for $<$, a true statement about probabilities almost always turns into a true statement about ranking functions. (There are but a few known exceptions to this transformation; Spohn 1994 mentions one.)

Two complaints about Jeffrey conditionalization carry over to Spohn conditionalization: Jeffrey respectively Spohn conditionalization is not commutative (Levi 1967b); any two regular probability measures respectively ranking functions can be related to each other via Jeffrey respectively Spohn conditionalization (by letting the evidential partition consist of the set of singletons containing the possibilities in W). The first complaint is misconceived, because both Jeffrey and Spohn conditionalization are *result-* rather than *evidence-oriented*: the parameter p_i respectively n_i characterizes the resulting degree of (dis)belief in E_i rather than the amount by which the evidence received between t and t' boosts or lowers the degree of (dis)belief in E_i. These parameters thus depend on both the prior epistemic state and the evidence received. Evidence first shifting E from p at t to p^* at t' and then to p'' at t^{**} is *not* a rearrangement of evidence first shifting E from p at t to p^{**} at t' and then to p^* at t''. Field (1978) respectively Shenoy (1991) presents a probabilistic respectively rank-theoretic update rule that is evidence-oriented in the sense of characterizing the evidence as such, independently of the prior epistemic state. Both of these rules are commutative.

The second complaint confuses input and output: Jeffrey respectively Spohn conditionalization does not rule out any evidential input as impossible (just as it does not rule out any prior epistemic state as impossible that is not already ruled out by the probability respectively ranking calculus). However, that does not imply that it is empty as a normative rule. On the contrary, for each prior epistemic and each evidential input there is one and only one posterior epistemic state that is compatible with Jeffrey respectively Spohn conditionalization. It is up to the agent what to do with a given epistemic state and a given evidential input, but it is up to nature which evidential input the agent receives.

One reason why an epistemic agent's degrees of belief should obey the probability calculus is that otherwise she is vulnerable to a Dutch Book (standard version) or an inconsistent evaluation of the fairness of bets (depragmatized version). For similar reasons she should update her subjective probability according to strict or Jeffrey conditionalization, depending on the format of the new infor-

mation. Why should degrees of disbelief obey the ranking calculus? And why should an epistemic agent update her ranking function according to plain or Spohn conditionalization?

The answers to these questions require a bit of terminology. An epistemic agent's degree of *entrenchment* for a proposition A is the number of "independent and minimally positively reliable" information sources saying A that it takes for the agent to give up her disbelief that A. If the agent does not disbelieve A to begin with, her degree of entrenchment for A is 0. If no finite number of information sources is able to make the agent give up her belief that A is false, her degree of entrenchment for A is ∞.

Suppose we want to determine Sophia's degree of entrenchment for the proposition that Vienna is the capital of Austria. This can be done by, say, putting her on the Stephansplatz and by counting the number of people passing by and telling her that Vienna is the capital of Austria. Her degree of entrenchment for the proposition that Vienna is the capital of Austria equals n precisely if she stops disbelieving that Vienna is the capital of Austria after n people have passed by and told her it is. The relation between these operationally defined degrees of entrenchment and the theoretical degrees of disbelief is similar to the relation between betting ratios and degrees of belief: under suitable conditions (when the information sources are independent and minimally positively reliable) the former can be used to measure the latter. Most of the time the conditions are not suitable, though. In Section 3.1 primitivism seemed to be the only plausible game in town. In the present case "going hypothetical" (Eriksson and Hájek 2007) is more promising: the agent's degree of disbelief in A is the number of information sources saying A that it would take for her to give up her disbelief that A, if those sources were independent and minimally positively reliable.

Now we are in the position to say why degrees of disbelief should obey the ranking calculus. They should do so, because an agent's belief set is and will always be consistent and deductively closed in the finite/countable/complete sense just in case her entrenchment function is a finitely/countably/completely minimitive ranking function and, depending on the format of the evidence, the agent updates according to plain or Spohn conditionalization (Huber 2007b).

It follows that the above definition of conditional ranks is the only good notion: both plain and Spohn conditionalization depend on the notion of conditional ranks, and the theorem does not hold if we replace that notion by another one. Furthermore, the definition of degrees of entrenchment makes only sense for natural numbers – after all, we have to *count* the independent and minimally positively reliable information sources. Therefore every concession to possibility theory – be it by adopting a different notion of conditional ranks or by allowing real-valued ranking functions – is a concession too much.

With the possible exception of decision making (see, however, Giang and Shenoy 2000), we can do everything with ranking functions that we can do with probability measures. In fact, in contrast to probability theory, ranking theory also has a notion of yes-or-no belief that is crucial if we want to stay in tune with traditional epistemology. (In addition, this allows for rank-theoretic theories of belief revision and of

nonmonotonic reasoning that are the topic of the next section.) Let me conclude this section with what I take to be a further advantage of ranking over probability theory.

Contrary to a widely held view there is no such thing as a genuinely unbiased assignment of probabilities (an *ur-* or *tabula rasa* prior, as we may call it) – even if we consider just a finite set of (more than two) possibilities. For instance, it is often said that assigning a probability of 1/6 to each of the six outcomes of a throw of a die is such an unbiased assignment. To see that this is not so it suffices to note that it follows from this assignment that the proposition that the number of spots the die will show after being thrown is greater than one is five times the probability of its negation. More generally, for every probability measure Pr on the powerset of $\{1, \ldots, 6\}$ there exists a contingent proposition A such that $\Pr(A) > \Pr(\overline{A})$. That is the sense in which there is no genuinely unbiased ur- or tabula rasa prior. This is in contrast to ranking theory, where the ur- or tabula rasa prior is that function $\varrho : \mathcal{A} \rightarrow N \cup \{\infty\}$ such that $\varrho(A) = 0$ for all consistent propositions A in $\mathcal{A}$, no matter how rich the field of propositions $\mathcal{A}$.

In probability theory we cannot adequately model conceptual changes – especially those that are due the agent's not being logically omniscient. Prior to learning a new concept the probabilistic agent is equipped with a probability measure Pr on some field $\mathcal{A}$ over some set W. When the agent learns a new concept, the possibilities ω in W become more fine grained. For instance, Sophia's set of enological possibilities with regard to a particular bottle of wine prior to learning the concept BARRIQUE is $W_1 = \{\text{red, white}\}$. After learning that concept her set of possibilities is

$$W_2 = \{\text{red \& barrique, red \& } \neg \text{ barrique, white \& barrique, white \& } \neg \text{ barrique}\}.$$

To model this conceptual change adequately, the field of propositions over the new set of possibilities W_2 will contain a counterpart-proposition for each old proposition in the field over W_1. In our example, the fields are the powersets. The counterpart-proposition of the old proposition that the bottle of wine is red, $\{\text{red}\} \subseteq W_1$, is $\{\text{red \& barrique, red \& } \neg \text{ barrique}\} \subseteq W_2$. The important epistemological feature of these conceptual changes is that Sophia does not learn any factual information; that is, she does not learn anything about which of the possibilities corresponds to the actual world. If ϱ_1 is Sophia's ranking function on the powerset of W_1, we want her ϱ_2 to be such that $\varrho_1(A) = \varrho_2(A')$ for each old proposition A in the powerset of W_1 and its counterpart proposition A' in the powerset of W_2, and such that $\varrho_2(B) = \varrho_2(\overline{B})$ for each (contingent) new proposition B. This is easily achieved by letting ϱ_2 copy ϱ_1 on the counterpart-propositions of the old propositions, and letting it copy the ur-prior on all the new propositions. In contrast to this there is no way of obtaining probability measures $\Pr_1$ on the old field and $\Pr_2$ on the new field that are related in this way.

The same is true for the different conceptual change that occurs when Sophia learns the new concept ROSÉ and thus that her old set of possibilities was not exclusive. If ϱ_1 is Sophia's ranking function on the powerset of W_1, her ϱ_3 on the powerset of $W_3 = \{\text{red, rosé, white}\}$ is that function ϱ_3 such that $\varrho_1(\{\omega\}) = \varrho_3(\{\omega\})$

for each old singleton-proposition $\{\omega\}$, and $\varrho_3\left(\{\omega'\}\right) = 0$ for each new singleton-proposition $\{\omega'\}$. Again, we cannot model this change in probability theory, since the only new probability measure that, in this sense, *conservatively* extends the old one assigns 0 to the (union of all) new possibilities. (Note that one and the same sentence may pick out different propositions with respect to the two sets of possibilities. For instance, with respect to W_1 the sentence "It is not a bottle of red wine" picks out the proposition that it is a bottle of white wine, {white}, while with respect to W_2 this sentence picks out the proposition that it is a bottle of rosé or white wine, {rosé, white}.) Arntzenius (2003) relies on just this inability of probability theory to cope with changes of the underlying set of worlds when he uses "spreading" to argue against conditionalization and reflection.

5 Belief Revision and Nonmonotonic Reasoning

5.1 Belief and Belief Revision

We have moved from degrees of belief to belief, and found ranking theory to provide a link between these two notions, thus satisfying the Lockean thesis. While some philosophers (most probabilists) hold the view that degrees of belief are more basic than beliefs, others adopt the opposite view. This is generally true of traditional epistemology, which is mainly concerned with the notion of knowledge and its tripartite definition as justified true belief. Belief in this sense comes in three "degrees": the ideally rational epistemic agent either believes A, or else she believes $W \setminus A$ and thus disbelieves A, or else she neither believes A nor $W \setminus A$ and thus suspends judgment with respect to A. Ordinary epistemic agents sometimes believe both A and $W \setminus A$, but since they should not do so, we may ignore this case.

According to this view an agent's epistemic state is characterized by the set of propositions she believes, her *belief set*. Such a belief set is required to be consistent and deductively closed (Hintikka 1961). In belief revision theory a belief set is usually represented as a set of sentences from a formal language $\mathcal{L}$ rather than as a set of propositions. The question addressed by belief revision theory (Alchourrón et al. 1985, Gärdenfors 1988, Gärdenfors and Rott 1995) is how an ideally rational epistemic agent should revise her belief set $\mathcal{B} \subseteq \mathcal{L}$ if she learns new information in form of a sentence α from $\mathcal{L}$. If α is consistent with $\mathcal{B}$ in the sense that $\mathcal{B} \nvdash \neg\alpha$, the agent should simply add α to $\mathcal{B}$ and close this set under (classical) logical consequence. In this case her new belief set, i.e. her old belief set $\mathcal{B}$ revised by the new information α, $\mathcal{B}\dotplus\alpha$, is the set of logical consequences of $\mathcal{B} \cup \{\alpha\}$:

$$\mathcal{B}\dotplus\alpha = Cn\left(\mathcal{B} \cup \alpha\right) = \{\beta \in \mathcal{L} : \mathcal{B} \cup \{\alpha\} \vdash \beta\}$$

Things get interesting when the new information α contradicts the old belief set $\mathcal{B}$. The basic idea is that the agent's new belief set $\mathcal{B}\dotplus\alpha$ should contain the new information α and as many of the old beliefs in $\mathcal{B}$ as is allowed by the requirement

that the new belief set be consistent and deductively closed. To state this more precisely, let us introduce the notion of a *contraction*. To contract a statement α from a belief set $\mathcal{B}$ is to give up the belief that α is true, but to keep as many of the remaining beliefs from $\mathcal{B}$ while ensuring consistency and deductive closure. Where $\mathcal{B} \dot{-} \alpha$ is the agent's new belief set after contracting her old belief set $\mathcal{B}$ by α, the A(lchourrón)G(ärdenfors)M(akinson) postulates for contraction $\dot{-}$ can be stated as follows. (Note that $\dot{+}$ as well as $\dot{-}$ are functions from $\wp(\mathcal{L}) \times \mathcal{L}$ into $\wp(\mathcal{L})$.) For every set of sentences $\mathcal{B} \subseteq \mathcal{L}$ and any sentences α and β in $\mathcal{L}$:

$\dot{-}1.$	If $\mathcal{B} = Cn(\mathcal{B})$, then $\mathcal{B} \dot{-} \alpha = Cn(\mathcal{B} \dot{-} \alpha)$.	Deductive Closure
$\dot{-}2.$	$\mathcal{B} \dot{-} \alpha \subseteq \mathcal{B}$.	Inclusion
$\dot{-}3.$	If $\alpha \notin Cn(\mathcal{B})$, then $\mathcal{B} \dot{-} \alpha = \mathcal{B}$.	Vacuity
$\dot{-}4.$	If $\alpha \notin Cn(\emptyset)$, then $\alpha \notin Cn(\mathcal{B} \dot{-} \alpha)$.	Success
$\dot{-}5.$	If $Cn(\{\alpha\}) = Cn(\{\beta\})$, then $\mathcal{B} \dot{-} \alpha = \mathcal{B} \dot{-} \beta$.	Preservation
$\dot{-}6.$	If $\mathcal{B} = Cn(\mathcal{B})$, then $\mathcal{B} \subseteq Cn((\mathcal{B} \dot{-} \alpha) \cup \{\alpha\})$.	Recovery
$\dot{-}7.$	If $\mathcal{B} = Cn(\mathcal{B})$, then $(\mathcal{B} \dot{-} \alpha) \cap (\mathcal{B} \dot{-} \beta) \subseteq \mathcal{B} \dot{-} (\alpha \wedge \beta)$.	
$\dot{-}8.$	If $\mathcal{B} = Cn(\mathcal{B})$ and $\alpha \notin \mathcal{B} \dot{-} (\alpha \wedge \beta)$, then $\mathcal{B} \dot{-} (\alpha \wedge \beta) \subseteq \mathcal{B} \dot{-} \alpha$.	

$\dot{-}1$ says that the contraction of $\mathcal{B}$ by α, $\mathcal{B} \dot{-} \alpha$, should be deductively closed, if $\mathcal{B}$ is deductively closed. $\dot{-}2$ says that a contraction should not give rise to new beliefs not previously held. $\dot{-}3$ says that the epistemic agent should not change her old beliefs when she gives up a sentence she does not believe to begin with. $\dot{-}4$ says that, unless α is tautological, the agent should really give up her belief that α is true if she contracts by α. $\dot{-}5$ says that the particular formulation of the sentence the agent gives up should not matter; in other words, the objects of belief shoud be propositions rather than sentences. $\dot{-}6$ says that the agent should recover her old beliefs if she first contracts by α and then adds α again. According to $\dot{-}7$ the agent should not give up more beliefs when contracting by $\alpha \wedge \beta$ than the ones she gives up when she contracts by α alone or by β alone. $\dot{-}8$ finally requires the agent not to give up more beliefs than necessary: if the agent gives up α when she contracts by $\alpha \wedge \beta$, she should not give up more than she gives up when contracting by α alone.

Given the notion of a contraction we can now state what the agent's new belief set $\mathcal{B} \dot{+} \alpha$ should look like. First, the agent should clear $\mathcal{B}$ to make it consistent with α. That is, the agent first should contract $\mathcal{B}$ by $\neg\alpha$. Then she should simply add α and close under (classical) logical consequence. The recipe just described is known as the *Levi identity*:

$$\mathcal{B} \dot{+} \alpha = Cn((\mathcal{B} \dot{-} \neg\alpha) \cup \{\alpha\})$$

Revision $\dot{+}$ defined in this way satisfies a corresponding list of properties. For every set of sentences $\mathcal{B} \subseteq \mathcal{L}$ and any sentences α and β in $\mathcal{L}$:

$\dot{+}1.$ $\mathcal{B} \dot{+} \alpha = Cn(\mathcal{B} \dot{+} \alpha)$.
$\dot{+}2.$ $\alpha \in \mathcal{B} \dot{+} \alpha$.
$\dot{+}3.$ If $\neg\alpha \notin Cn(\mathcal{B})$, then $\mathcal{B} \dot{+} \alpha = Cn(\mathcal{B} \cup \{\alpha\})$.
$\dot{+}4.$ If $\neg\alpha \notin Cn(\emptyset)$, then $\bot \notin \mathcal{B} \dot{+} \alpha$.

$\dotplus$5. If $Cn(\{\alpha\}) = Cn(\{\beta\})$, then $\mathcal{B}\dotplus\alpha = \mathcal{B}\dotplus\beta$.

$\dotplus$6. If $\mathcal{B} = Cn(\mathcal{B})$, then $(\mathcal{B}\dotplus\alpha) \cap \mathcal{B} = \mathcal{B}\dotdiv\neg\alpha$.

$\dotplus$7. If $\mathcal{B} = Cn(\mathcal{B})$, then $\mathcal{B}\dotplus(\alpha \wedge \beta) \subseteq Cn((\mathcal{B}\dotplus\alpha) \cup \{\beta\})$.

$\dotplus$8. If $\mathcal{B} = Cn(\mathcal{B})$ and $\neg\beta \notin \mathcal{B}\dotplus\alpha$, then $Cn((\mathcal{B}\dotplus\alpha) \cup \{\beta\}) \subseteq \mathcal{B}\dotplus(\alpha \wedge \beta)$.

(The contradictory sentence $\bot$ can be defined as the negation of the tautological sentence $\top$, $\neg\top$.) Rott (2001) discusses many further principles and variations of the above.

In standard belief revision theory the new information is always part of the new belief set. *Non-prioritized* belief revision relaxes this requirement (Hansson 1999). Personally I am not quite sure this makes sense conceptually, but the idea seems to be that the agent might add the new information and then check for consistency, which makes her give up part or all of the new information again, because her old beliefs turn out to be more *entrenched*. (The degrees of entrenchment mentioned in the previous section are named after this relation, but it is to be noted that the former are operationally defined, while the latter is a theoretical notion).

The notion of entrenchment provides the connection to degrees of belief. In order to decide which part of her belief set she wants to give up, belief revision theory equips the ideally rational epistemic agent with an *entrenchment ordering*. Technically, this is a relation $\preceq$ on $\mathcal{L}$ such that for all α, β, and γ in $\mathcal{L}$:

E1. If $\alpha \preceq \beta$ and $\beta \preceq \gamma$, then $\alpha \preceq \gamma$.   Transitivity

E2. If $\alpha \vdash \beta$, then $\alpha \preceq \beta$.   Dominance

E3. $\alpha \preceq \alpha \wedge \beta$ or $\beta \preceq \alpha \wedge \beta$.   Conjunctivity

E4. If $\bot \notin Cn(\mathcal{B})$, then $\alpha \notin \mathcal{B}$ just in case $\forall\beta \in \mathcal{L} : \alpha \preceq \beta$.   Minimality

E5. If $\forall\alpha \in \mathcal{L} : \alpha \preceq \beta$, then $\beta \in Cn(\emptyset)$.   Maximality

$\mathcal{B}$ is a fixed set of background beliefs. Given an entrenchment ordering $\preceq$ on $\mathcal{L}$ we can define a revision $\dotplus$ as follows:

$$\mathcal{B}\dotplus\alpha = \{\beta \in \mathcal{B} : \neg\alpha \prec \beta\} \cup \{\alpha\}$$

Here $\alpha \prec \beta$ holds just in case $\alpha \preceq \beta$ and $\beta \npreceq \alpha$. Then one can prove the following *representation theorem*.

Theorem 1 *Let $\mathcal{L}$ be a formal language, let $\mathcal{B} \subseteq \mathcal{L}$ be a set of sentences, and let α be a sentence in $\mathcal{L}$. Each entrenchment ordering $\preceq$ on $\mathcal{L}$ induces a revision operator $\dotplus$ on $\mathcal{L}$ satisfying $\dotplus1$-$\dotplus8$ by defining $\mathcal{B}\dotplus\alpha = \{\beta \in \mathcal{B} : \neg\alpha \prec \beta\} \cup \{\alpha\}$. For each revision operator $\dotplus$ on $\mathcal{L}$ satisfying $\dotplus1$-$\dotplus8$ there is an entrenchment ordering $\preceq$ on $\mathcal{L}$ that induces $\dotplus$ in exactly this way.*

It is, however, fair to say that belief revision theorists distinguish between degrees of belief and entrenchment. Entrenchment, so they say, characterizes the agent's unwillingness to give up a particular belief, which may be different from her degree of belief for the respective sentence or proposition. Although this distinction seems to violate Occam's razor by unnecessarily introducing an additional epistemic level, it corresponds to Spohn's parallelism (see Spohn's contribution to this

volume) between subjective probabilities and ranking functions as well as to Stalnaker's stance in his (1996: sct. 3). Weisberg (to appear: sct. 1.7) expresses similar sentiments.

Suppose the agent's epistemic state is represented by a ranking function ϱ (on a field of propositions over the set of models $Mod_{\mathcal{L}}$ for the language $\mathcal{L}$, as explained in Section 1). Then the ordering $\preceq_{\varrho}$ that is defined for all α and β in $\mathcal{L}$ by

$$\alpha \preceq_{\varrho} \beta \quad \text{if and only if} \quad \varrho\left(Mod\left(\neg\alpha\right)\right) \leq \varrho\left(Mod\left(\neg\beta\right)\right)$$

is an entrenchment ordering for $\mathcal{B} = \{\alpha \in \mathcal{L} : \varrho\left(Mod\left(\neg\alpha\right)\right) > 0\}$.

Ranking theory thus covers AGM belief revision theory as a special case. It is important to see how ranking theory goes beyond AGM belief revision theory. In the latter theory the agent's prior epistemic state is characterized by a belief set $\mathcal{B}$ together with an entrenchment ordering $\preceq$. If the agent receives new information in form of a sentence α, the entrenchment ordering is used to turn the old belief set into new one, viz. $\mathcal{B} \dotplus \alpha$. The agent's posterior epistemic state is thus characterized by a belief set only. The entrenchment ordering itself is not updated. Therefore AGM belief revision theory cannot handle iterated belief changes. To the extent that belief revision is not simply a one shot game, AGM belief revision theory is thus no theory of belief revision at all. (The analogous situation in terms of subjective probabilities is to characterize the agent's prior epistemic state by a set of propositions together with a subjective probability measure, and to use that measure to update the set of propositions without ever changing the probability measure itself.)

In ranking theory the agent's prior epistemic state is characterized by a ranking function ϱ (on a field over $Mod_{\mathcal{L}}$). That function determines the agent's prior belief set $\mathcal{B}$, and so there is no need to additionally specify $\mathcal{B}$. If the agent receives new information in form of a proposition A, as (the propositional equivalent of) AGM belief revision theory has it, there are infinitely many ways to update her ranking function that all give rise to the same new belief set $\mathcal{B} \dotplus A$. Let n be an arbitrary positive number in $N \cup \{\infty\}$. Then Spohn conditionalization on the partition $\{A, Mod_{\mathcal{L}} \setminus A\}$ with $n > 0$ as new rank for $Mod_{\mathcal{L}} \setminus A$ (and consequently 0 as new rank for A), $\varrho'_n\left(Mod_{\mathcal{L}} \setminus A\right) = n$, determines a new ranking function ϱ'_n that induces a belief set $\mathcal{B}'_n$. It holds for any two positive numbers m and n in $N \cup \{\infty\}$:

$$\mathcal{B}'_m = \mathcal{B}'_n = \mathcal{B} \dotplus A,$$

where the latter is the belief set described two paragraphs ago.

Plain conditionalization is the special case of Spohn conditionalization with ∞ as new rank for $Mod_{\mathcal{L}} \setminus A$. The new ranking function obtained in this way is $\varrho'_{\infty} = \varrho(\cdot \mid A)$, and the belief set it induces is the same $\mathcal{B} \dotplus A$ as before. However, once the epistemic agent assigns rank ∞ to $Mod_{\mathcal{L}} \setminus A$, she can never get rid of A again (in the sense that the only information that would allow her to give up her belief that A is to become certain that A is false, i.e. assign rank ∞ to A; that in turn would make her epistemic state collapse in the sense of turning it into the tabula rasa ranking from Section 4 that is agnostic with respect to all contingent

propositions). Just as one is stuck with A once one assigns it probability 1, so one is basically stuck with A once one assigns its negation rank ∞. As we have seen, AGM belief revision theory is compatible with always updating in this way. That explains why it cannot handle iterated belief revision. To rule out this behavior one has to impose further constraints on entrenchment orderings. Boutilier (1996) as well as Darwiche and Pearl (1997) do so by postulating constraints compatible with, but not yet implying ranking theory. Hild and Spohn (2008) argue that one really has to go all the way to ranking theory. Stalnaker (2009) critically discusses these approaches and argues that one needs to distinguish different kinds of information, including meta-information about the agent's own beliefs and revision policies as well as about the sources of her information.

5.2 Belief and Nonmonotonic Reasoning

A premise β classically entails a conclusion γ, $\beta \vdash \gamma$, just in case γ is true in every model or truth value assignment in which β is true. The classical consequence relation $\vdash$ (conceived of as a relation between two sentences, i.e. $\vdash \subseteq \mathcal{L} \times \mathcal{L}$, rather than as a relation between a set of sentences, the premises, and a sentence, the conclusion) is *non-ampliative* in the sense that the conclusion of a classically valid argument does not convey information that goes beyond the information contained in the premise.

$\vdash$ has the following *monotonicity* property. For any sentences α, β, and γ in $\mathcal{L}$:

$$\text{If } \beta \vdash \gamma, \text{ then } \alpha \wedge \beta \vdash \gamma.$$

That is, if γ follows from β, then γ follows from any sentence $\alpha \wedge \beta$ that is at least as logically strong as β. However, everyday reasoning is often ampliative. When Sophia sees the thermometer at 33° Celsius she infers that it is not too cold to wear her sundress. If Sophia additionally sees that the thermometer is placed above the oven where she is boiling her pasta, she will not infer that anymore. *Nonmonotonic reasoning* is the study of reasonable consequence relations which violate monotonicity (Gabbay 1985, Makinson 1989, Kraus et al. 1990; for an overview see Makinson 1994).

For a fixed set of background beliefs $\mathcal{B}$, the revision operators $\dot{+}$ from the previous section give rise to nonmonotonic consequence relations $\vert\sim$ as follows (Makinson and Gärdenfors 1991):

$$\alpha \vert\sim \beta \quad \text{if and only if} \quad \beta \in \mathcal{B}\dot{+}\alpha$$

Nonmonotonic consequence relations on a language $\mathcal{L}$ are supposed to satisfy the following principles from Kraus et al. (1990).

KLM1. $\alpha \vert\sim \alpha$.   Reflexivity
KLM2. If $\vdash \alpha \leftrightarrow \beta$ and $\alpha \vert\sim \gamma$, then $\beta \vert\sim \gamma$.   Left Logical Equivalence
KLM3. If $\vdash \alpha \rightarrow \beta$ and $\gamma \vert\sim \alpha$, then $\gamma \vert\sim \beta$.   Right Weakening

KLM4. If $\alpha \wedge \beta \mathrel{|\!\sim} \gamma$ and $\alpha \mathrel{|\!\sim} \beta$, then $\alpha \mathrel{|\!\sim} \gamma$.   Cut

KLM5. If $\alpha \mathrel{|\!\sim} \beta$ and $\alpha \mathrel{|\!\sim} \gamma$, then $\alpha \wedge \beta \mathrel{|\!\sim} \gamma$.   Cautious Monotonicity

KLM6. If $\alpha \mathrel{|\!\sim} \beta$ and $\alpha \mathrel{|\!\sim} \gamma$, then $\alpha \vee \beta \mathrel{|\!\sim} \gamma$.   Or

The standard interpretation of a nonmonotonic consequence relation $\mathrel{|\!\sim}$ is "If ..., normally ...". Normality among worlds is spelt out in terms of *preferential models* $\langle S, l, \prec \rangle$ for $\mathcal{L}$, where S is a set of states and $l : S \to Mod_{\mathcal{L}}$ is a function from S to the set of models for $\mathcal{L}$, $Mod_{\mathcal{L}}$, that assigns each state s its model or world $l(s)$. The abnormality relation $\prec$ is a strict partial order on $Mod_{\mathcal{L}}$ that satisfies a certain "smoothness" condition. For our purposes it suffices to note that the order among the worlds that is induced by a pointwise ranking functions is such an abnormality relation. Given a preferential model $\langle S, l, \prec \rangle$ we can define a nonmonotonic consequence relation $\mathrel{|\!\sim}$ as follows. Let $\widehat{\alpha}$ be the set of states in whose worlds α is true, i.e. $\widehat{\alpha} = \{s \in S : l(s) \models \alpha\}$, and define

$$\alpha \mathrel{|\!\sim} \beta \quad \text{if and only if} \quad \forall s \in \widehat{\alpha} : \text{if } \forall t \in \widehat{\alpha} : t \not\prec s, \text{ then } l(s) \models \beta.$$

That is, $\alpha \mathrel{|\!\sim} \beta$ holds just in case β is true in the least abnormal among the α-worlds. Then one can prove the following representation theorem.

Theorem 2 *Let $\mathcal{L}$ be a language, let $\mathcal{B} \subseteq \mathcal{L}$ be a set of sentences, and let α be a sentence in $\mathcal{L}$. Each preferential model $\langle S, l, \prec \rangle$ for $\mathcal{L}$ induces a nonmonotonic consequence relation $\mathrel{|\!\sim}$ on $\mathcal{L}$ satisfying KLM1-KLM6 by defining*

$$\alpha \mathrel{|\!\sim} \beta \text{ if and only if } \forall s \in \widehat{\alpha} : \text{if } \forall t \in \widehat{\alpha} : t \not\prec s, \text{ then } l(s) \models \beta.$$

For each nonmonotonic consequence relation on $\mathcal{L}$ satisfying KLM1-KLM6 there is a preferential model $\langle S, l, \prec \rangle$ for $\mathcal{L}$ that induces $\mathrel{|\!\sim}$ in exactly this way.

Whereas the classical consequence relation preserves truth in all logically possible worlds, nonmonotonic consequence relations preserve truth in all least abnormal worlds. For a different semantics in terms of inhibition nets see Leitgeb (2004).

What is of particular interest to us is the fact that these nonmonotonic consequence relations can be induced by a fixed set of background beliefs $\mathcal{B}$ and various forms of degrees of belief over $\mathcal{B}$. We will not attempt to indicate how this works. Makinson's contribution to this volume is an excellent presentation of ideas underlying nonmonotonic reasoning and its relation to degrees of belief. Similar remarks apply to Rott's contribution to this volume, in which entrenchment orderings, ranking functions, and further models of epistemic states are defined for beliefs as well as disbeliefs and non-beliefs.

References

Alchourrón, Carlos E., and Gärdenfors, Peter and Makinson, David (1985), On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic* **50**, 510–530.

Armendt, Brad (1980), Is There a Dutch Book Argument for Probability Kinematics? *Philosophy of Science* **47**, 583–588.

Arntzenius, Frank (2003), Some Problems for Conditionalization and Reflection. *Journal of Philosophy* **100**, 356–371.

Boutilier, Craig (1996), Iterated Revision and Minimal Change of Belief. *Journal of Philosophical Logic* **25**, 263–305.

Bronfman, Aaron (manuscript), A Gap in Joyce's Argument for Probabilism. University of Michigan: unpublished manuscript.

Carnap, Rudolf (1962), *Logical Foundations of Probability*. 2nd ed. Chicago: University of Chicago Press.

Christensen, David (2004), *Putting Logic in Its Place. Formal Constraints on Rational Belief*. Oxford: Oxford University Press.

Cox, Richard T. (1946), Probability, Frequency, and Reasonable Expectation. *American Journal of Physics* **14**, 1–13.

Darwiche, Adnan and Pearl, Judea (1997), On the Logic of Iterated Belief Revision. *Artificial Intelligence* **89**, 1–29.

Dempster, Arthur P. (1968), A Generalization of Bayesian Inference. *Journal of the Royal Statistical Society. Series B (Methodological)* **30**, 205–247.

Dubois, Didier and Prade, Henri (1988), *Possibility Theory. An Approach to Computerized Processing of Uncertainty*. New York: Plenum.

Egan, Andy (2006), Secondary Qualities and Self-Location. *Philosophy and Phenomenological Research* **72**, 97–119.

Elga, Adam (2000), Self-Locating Belief and the Sleeping Beauty Problem. *Analysis* **60**, 143–147.

Eriksson, Lina and Hájek, Alan (2007), What Are Degrees of Belief? *Studia Logica* **86**, 183–213.

Field, Hartry (1978), A Note on Jeffrey Conditionalization. *Philosophy of Science* **45**, 361–367.

Foley, Richard (1992), The Epistemology of Belief and the Epistemology od Degrees of Belief. *American Philosophical Quaterly* **29**, 111–121.

Frankish, Keith (2004), *Mind and Supermind*. Cambridge: Cambridge University Press.

Gabbay, Dov M. (1985), Theoretical Foundations for Non-Monotonic Reasoning in Expert Systems. In K.R. Apt (ed.), *Logics and Models of Concurrent Systems*. NATO ASI Series **13**. Berlin: Springer, 439–457.

Gärdenfors, Peter (1988), *Knowledge in Flux. Modeling the Dynamics of Epistemic States*. Cambridge, MA: MIT Press.

Gärdenfors, Peter and Rott, Hans (1995), Belief Revision. In D.M. Gabbay, and C.J. Hogger and J.A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming. Vol. 4: Epistemic and Temporal Reasoning*. Oxford: Clarendon Press, 35–132.

Giang, Phan H. and Shenoy, Prakash P. (2000), A Qualitative Linear Utility Theory for Spohn's Theory of Epistemic Beliefs. In C. Boutilier and M. Goldszmidt (eds.), *Uncertainty in Artificial Intelligence* **16**. San Francisco: Morgan Kaufmann, 220–229.

Haenni, Rolf and Lehmann, Norbert (2003), Probabilistic Argumentation Systems: A New Perspective on Dempster-Shafer Theory. *International Journal of Intelligent Systems* **18**, 93–106.

Hájek, Alan (2003), What Conditional Probability Could Not Be. *Synthese* **137**, 273–323.

Hájek, Alan (2007), Probability, Interpretations of. In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/probability-interpret/

Halpern, Joseph Y. (2003), *Reasoning About Uncertainty*. Cambridge, MA: MIT Press.

Hansson, Sven Ove (1999), A Survey of Non-Prioritized Belief Revision. *Erkenntnis* **50**, 413–427.

Hawthorne, James and Bovens, Luc (1999), The *Preface*, the *Lottery*, and the Logic of Belief. *Mind* **108**, 241–264.

Hempel, Carl Gustav (1962), Deductive-Nomological vs. Statistical Explanation. In H. Feigl and G. Maxwell (eds.), *Scientific Explanation, Space and Time. Minnesota Studies in the Philosophy of Science* **3**. Minneapolis: University of Minnesota Press, 98–169.

Hild, Matthias and Spohn, Wolfgang (2008), The Measurement of Ranks and the Laws of Iterated Contraction. *Artificial Intelligence* **172**, 1195–1218.

Hintikka, Jaakko (1961), *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Ithaca, NY: Cornell University Press. Reissued as J. Hintikka (2005), *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Prepared by V.F. Hendricks and J. Symons. London: College Publications.

Huber, Franz (2006), Ranking Functions and Rankings on Languages. *Artificial Intelligence* **170**, 462–471.

Huber, Franz (2007a), Confirmation and Induction. In J. Fieser & B. Dowdon (eds.), *The Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/c/conf-ind.htm

Huber, Franz (2007b), The Consistency Argument for Ranking Functions. *Studia Logica* **86**, 299–329.

Jeffrey, Richard C. (1970), Dracula Meets Wolfman: Acceptance vs. Partial Belief. In M. Swain (ed.), *Induction, Acceptance, and Rational Belief*. Dordrecht: Reidel, 157–185.

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

Jeffrey, Richard (2004), *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press.

Joyce, James M. (1998), A Nonpragmatic Vindication of Probabilism. *Philosophy of Science* **65**, 575–603.

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

Kahneman, Daniel, Slovic, Paul and Tversky, Amos, eds., (1982), *Judgment Under Uncertainty. Heuristics and Biases*. Cambridge: Cambridge University Press.

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

Kneale, William C. (1949), *Probability and Induction*. Oxford: Clarendon Press.

Kolmogorov, Andrej N. (1956), *Foundations of the Theory of Probability*. 2nd ed. New York: Chelsea Publishing Company.

Krantz, David H., Luce, Duncan R., Suppes, Patrick and Tversky, Amos (1971), *Foundations of Measurement*. Vol. I. New York: Academic Press.

Kraus, Sarit, Lehmann, Daniel and Magidor, Menachem (1990), Nonmonotonic Reasoning, Preferential Models, and Cumulative Logics. *Artificial Intelligence* **40**, 167–207.

Kripke, Saul (1979), A Puzzle About Belief. In A. Margalit (ed.), *Meaning and Use*. Dordrecht: D. Reidel, 239–283.

Kvanvig, Jonathan L. (1994), A Critique of van Fraassen's Voluntaristic Epistemology. *Synthese* **98**, 325–348.

Kyburg, Henry E. Jr. (1961), *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press.

Kyburg, Henry E. Jr. and Teng, Choh Man (2001), *Uncertain Inference*. Cambridge: Cambridge University Press.

Leitgeb, Hannes (2004), *Inference on the Low Level*. Dordrecht: Kluwer.

Levi, Isaac (1967a), *Gambling With Truth. An Essay on Induction and the Aims of Science*. New York: Knopf.

Levi, Isaac (1967b), Probability Kinematics. *British Journal for the Philosophy of Science* **18**, 197–209.

Levi, Isaac (1978), Dissonance and Consistency according to Shackle and Shafer. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association.* Vol. II: Symposia and Invited Papers, 466–477.

Levi, Isaac (1980), *The Enterprise of Knowledge.* Cambridge, MA: MIT Press.

Lewis, David K. (1979), Attitudes *De Dicto* and *De Se. The Philosophical Review* **88**, 513–543. Reprinted with postscripts in D. Lewis (1983), *Philosophical Papers.* Vol. I. Oxford: Oxford University Press, 133–159.

Lewis, David K. (1980), A Subjectivist's Guide to Objective Chance. In R.C. Jeffrey (ed.), *Studies in Inductive Logic and Probability.* Vol. II. Berkeley: University of Berkeley Press, 263–293. Reprinted in D. Lewis (1986), *Philosophical Papers.* Vol. II. Oxford: Oxford University Press, 83–113.

Lewis, David K. (1986), *On the Plurality of Worlds.* Oxford: Blackwell.

Lewis, David K. (1999), Why Conditionalize? In D. Lewis (1999), *Papers in Metaphysics and Epistemology.* Cambridge: Cambridge University Press, 403–407.

Lewis, David K. (2001), Sleeping Beauty: Reply to Elga. *Analysis* **61**, 171–176.

Locke, John (1690/1975), *An Essay Concerning Human Understanding.* Oxford: Clarendon Press.

Maher, Patrick (2006), Review of David Christensen, *Putting Logic in Its Place. Formal Constraints on Rational Belief. Notre Dame Journal of Formal Logic* **47**, 133–149.

Makinson, David (1965), The Paradox of the Preface. *Analysis* **25**, 205–207.

Makinson, David (1989), General Theory of Cumulative Inference. In M. Reinfrank, J. de Kleer, M.L. Ginsberg and E. Sandewall (eds.), *Non-Monotonic Reasoning.* Lecture Notes in Artificial Intelligence **346**. Berlin: Springer, 1–18.

Makinson, David (1994), General Patterns in Nonmonotonic Reasoning. In D.M. Gabbay, C.J. Hogger and J.A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming. Vol. 3: Nonmonotonic Reasoning and Uncertain Reasoning.* Oxford: Clarendon Press, 35–110.

Makinson, David and Gärdenfors, Peter (1991), Relations between the Logic of Theory Change and Nonmonotonic Logic. A. Fuhrmann and M. Morreau (eds.), *The Logic of Theory Change.* Berlin: Springer, 185–205.

Paris, Jeff B. (1994), *The Uncertain Reasoner's Companion – A Mathematical Perspective. Cambridge Tracts in Theoretical Computer Science* **39**. Cambridge: Cambridge University Press.

Ramsey, Frank P. (1926), Truth and Probability. In F.P. Ramsey (1931), *The Foundations of Mathematics and Other Logical Essays.* Ed. by R.B. Braithwaite. London: Kegan, Paul, Trench, Trubner & Co., New York: Harcourt, Brace and Company, 156–198.

Roorda, Jonathan (1995), Revenge of Wolfman: A Probabilistic Explication of Full Belief. http://www.princeton.edu/bayesway/pu/Wolfman.pdf

Rott, Hans (2001), *Change, Choice, and Inference. A Study of Belief Revision and Nonmonotonic Reasoning.* Oxford: Oxford University Press.

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

Shackle, George L.S. (1949), *Expectation in Economics.* Cambridge: Cambridge University Press.

Shackle, George L.S. (1969), *Decision, Order, and Time.* 2nd ed. Cambridge: Cambridge University Press.

Shafer, Glenn (1976), *A Mathematical Theory of Evidence.* Princteton, NJ: Princeton University Press.

Shenoy, Prakash P. (1991), On Spohn's Rule for Revision of Beliefs. *International Journal of Approximate Reasoning* **5**, 149–181.

Skyrms, Brian (1987), Dynamic Coherence and Probability Kinematics. *Philosophy of Science* **54**, 1–20.

Smets, Philippe (2002), Showing Why Measures of Quantified Beliefs are Belief Functions. In B. Bouchon, and L. Foulloy and R.R. Yager (eds.), *Intelligent Systems for Information Processing: From Representation to Applications.* Amsterdam: Elsevier, 265–276.

Smets, Philippe and Kennes, Robert (1994), The Transferable Belief Model. *Artifical Intelligence* **66**, 191–234.

Spohn, Wolfgang (1988), Ordinal Conditional Functions: A Dynamic Theory of Epistemic States. In W.L. Harper and B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics* **II**. Dordrecht: Kluwer, 105–134.

Spohn, Wolfgang (1990), A General Non-Probabilistic Theory of Inductive Reasoning. In R.D. Shachter, T.S. Levitt, J. Lemmer and L.N. Kanal (eds.), *Uncertainty in Artificial Intelligence* **4**. Amsterdam: North-Holland, 149–158.

Spohn, Wolfgang (1994), On the Properties of Conditional Independence. In P. Humphreys (ed.), *Patrick Suppes. Scientific Philosopher. Vol. 1: Probability and Probabilistic Causality*. Dordrecht: Kluwer, 173–194.

Stalnaker, Robert C. (1996), Knowledge, Belief and Counterfactual Reasoning in Games. *Economics and Philosophy* **12**, 133–162.

Stalnaker, Robert C. (2003), *Ways a World Might Be*. Oxford: Oxford University Press.

Stalnaker, Robert C. (2009), Iterated Belief Revision. *Erkenntnis* **70** (1).

Steup, Matthias (2006), Knowledge, Analysis of. In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/knowledge-analysis/

Teller, Paul (1973), Conditionalization and Observation. *Synthese* **26**, 218–258.

van Fraassen, Bas C. (1989), *Laws and Symmetry*. Oxford: Oxford University Press.

van Fraassen, Bas C. (1990), Figures in a Probability Landscape. In J.M. Dunn and A. Gupta (eds.), *Truth or Consequences*. Dordrecht: Kluwer, 345–356.

van Fraassen, Bas C. (1995), Belief and the Problem of Ulysses and the Sirens. *Philosophical Studies* **77**, 7–37.

Walley, Peter (1991), *Statistical Reasoning With Imprecise Probabilities*. New York: Chapman and Hall.

Weisberg, Jonathan (to appear), Varieties of Bayesianism. In D.M. Gabbay, S. Hartmann and J. Woods (eds.), *Handbook of the History of Logic. Vol. 10: Inductive Logic*. Amsterdam/New York: Elsevier.

Zadeh, Lotfi A. (1978), Fuzzy Sets as a Basis for a Theory of Possibility. *Fuzzy Sets and Systems* **1**, 3–28.