9

FULL & PARTIAL BELIEF

Konstantin Genin

Philosophers and scientists in allied fields use the term 'belief' to refer roughly to the attitude taken toward a proposition regarded as true. That first approximation is unlikely to satisfy those in search of a non-circular definition. Early twentieth-century psychologists and philosophers of mind attempted to address that difficulty by reducing belief to some sort of behavioral disposition. Although the behaviorist project is usually taken to have been a failure, there is broad consensus that belief leaves a distinctive behavioral footprint: most philosophers would agree that an agent who believes P can be expected to accept P as a premise in reasoning, planning and deliberation. If I believe that the A train is running express, I will plan to take it if going to Harlem, but try to catch the local if going to the Museum of Natural History. Similarly banal examples can be easily multiplied.

It is commonly held that belief is not just an all-or-nothing matter, but admits of degrees. A veteran subway rider may have a higher degree of belief in the proposition that the A train will run local next weekend than in the proposition that the A train will run local next rush hour. Philosophers sufficiently impressed by examples of this sort orient their activity around the structure of "partial belief" rather than the all-or-nothing attitude denoted by "full belief," or belief *simpliciter.* Although it is easy to generate plausible examples of partial beliefs, it is harder to say exactly what is meant by a degree of belief. An agent's degree of belief in P may reflect their level of confidence in the truth of P, their willingness to assent to P in conversation, or perhaps how much evidence is required to convince them to abandon their belief in P. A venerable tradition, receiving classical expression in Ramsey (1931) and de Finetti (1937), holds that degrees of belief are most directly reflected in which bets regarding P an agent is willing to accept. At least since Pascal, mainstream philosophical opinion has held that degrees of belief are well-modeled by probabilities (see Hacking, 1975, for a readable history). To this day, subjective, or "epistemic," probability remains one of the dominant interpretations of the probability calculus.

A parallel tradition, though never as dominant, holds that degrees of belief are neither so precise, nor as definitely comparable as suggested by Pascal's probabilistic analysis. Keynes (1921) famously proposes that degrees of belief may enjoy only an *ordinal* structure, which admits of qualitative, but not quantitative, comparison. Keynes even suggests that the strength of some pairs of partial beliefs cannot be compared at all.

Cohen (1980) traces another minority tradition to Francis Bacon's *Novum Organum*. On the usual probability scale a degree of belief of zero in some proposition implies maximal conviction in its negation. On the Baconian scale, a degree of belief of zero implies no conviction in either the proposition or its negation. Thus, the usual scale runs from "disproof to proof" whereas the Baconian runs from "no evidence, or non-proof to proof" (Cohen, 1980, p. 224). In the past few decades, Baconian probability has received increasing attention, resulting in theories approaching the maturity and sophistication of those in the Pascalian tradition (Spohn, 2012; Huber, this volume).

Formal epistemologists are traditionally interested in both full and partial belief, although most would probably take partial belief as the primary object of study. Moss (2018) even argues that there are instances of probabilistic *knowledge* that do not involve any full beliefs. On the other hand, traditional analytic epistemologists and philosophers of mind routinely study full belief and related all-or-nothing attitudes such as knowledge and desire, but only rarely show interest in their graded counterparts. The differential emphasis on partial beliefs, although often commented upon, may reflect sociological factors more than any essential difference between the fields. These differences will likely become less pronounced in the future.

What is less often remarked upon is traditional epistemology's focus on individual beliefs, rather than entire systems of belief, as is typical in formal epistemology. Traditional philosophers are interested in what it means for an agent S to believe a particular proposition P. Representationalist philosopher of mind wonder how intentional states, or states that involve "aboutness" arise at all, especially if the agents involved are correctly understood as purely physical systems. Formal epistemologists tend to take matters of mental representation for granted, rarely inquiring into how the trick is worked. Dispositionalist philosophers of mind are interested in analyzing an agent's belief that P into a disposition to reason or act, although they will disagree about how readily these dispositions will be observed in behavior. Their focus on individual beliefs gives rise to certain standard objections. A Muscovite who believes, in the 1930s, that the Stalinist terror is morally wrong, may not betray her beliefs in her behavior at all.

Formal epistemologists resolve such difficulties by insisting on a holism about belief: it is entire *systems* of belief (and perhaps utility) that are reflected in deliberation and action, otherwise underdetermined by individual beliefs. In general, formal epistemologists are interested in the norms governing the structure and dynamics of whole systems of full or partial belief: how individual beliefs must systematically cohere in order to be rational; how they must be reflected in decision making; and how they ought to accommodate new evidence. Accordingly, those issues will be the focus of this article. For a good introduction to belief in the philosophy of mind, see Schwitzgebel (2015). See Hájek and Lin (2017) for a suggestive discussion of how mainstream and formal epistemology would benefit from increased sensitivity to each other's concerns.

Not everyone agrees that both partial and full beliefs exist—there are theorists who attempt to eliminate one or the other attitude. But anyone who admits the existence of both full and partial belief inherits a thorny problem: how are full beliefs related to partial beliefs? Two answers immediately suggest themselves. The first claims that full belief is just the maximal degree of partial belief. The second argues that full belief is just partial belief above a certain threshold. Both answers give rise to formidable problems. Other theorists claim that an agent's partial beliefs underdetermine their full beliefs in the absence of information about the agent's preferences.

In the last few years, the question of how partial and full belief are related has received considerable attention in formal epistemology, giving rise to several subtle, elegant and, unfortunately, incompatible solutions. The debate between these alternatives is the heart of this article and is presented in Section 5. The preceding sections develop the context and background necessary to understand and appreciate this debate. Readers who feel comfortable with these prerequisites, as well as those who are in a hurry, may skip to the final section and refer back to previous sections only as necessary.

1 THE OBJECTS OF BELIEF

In the following we will see several proposed models for the structure of belief. Most of these proposals take the objects of belief to be either *propositions*, or *sentences* in a formalized language. This section reviews the basic notions required to work with propositions and sentences. If the reader feels overwhelmed with the technicalities in this section, they should feel free to postpone them, and refer back to it on-the-fly. Readers who are accustomed to working with these objects may freely skip this section.

For our purposes, a *possible world* is a way the world, or some interesting aspect of the world, might be. We let W denote the set of *all* possible worlds, i.e. the set of all possible ways the world might be. It is not necessary to think of these as objective, metaphysical realities. More often, possible worlds are constrained by contextual presuppositions, and their granularity reflects our interests. Suffice it to say that knowing the *true* possible world $w \in W$ would satisfy an agent's curiosity—she would thereby settle some interesting matter under discussion. A proposition $P \subseteq W$ is a *set* of possible worlds, i.e. it is a partial specification of the way the world is. To know that P is true is to know that the true world is among the set of worlds $\{w : w \in P\}$ since P is true in a possible world w iff $w \in P$.

Propositions enjoy a set-theoretic structure. The relative complement of P, $\neg P = W \setminus P$, is the set of all worlds in which P is false. If P, Q are arbitrary propositions, then their intersection $P \cap Q$ is the set of all worlds in which P and Q are both true. The disjunction $P \cup Q$ is the set of worlds in which at least one of P, Q is true. The material conditional $P \to Q$ is the set of worlds $\neg P \cup Q$, in which either P is false or Q is true. If $P \subseteq Q$ we say that P *entails* Q and also that P is *logically stronger* than Q. If $P \subseteq Q$ and $Q \subseteq P$ we write $P \equiv Q$ and say that P and Q are *logically equivalent*. The tautological proposition W is true in all worlds and the contradictory proposition, the empty set $\varnothing$, is not true in any world. A set of propositions $\mathcal{A}$ is *consistent* iff there is a world in which all the elements of $\mathcal{A}$ are true, i.e. if $\cap\mathcal{A} \neq \varnothing$. Otherwise, we say that $\mathcal{A}$ is *inconsistent*. A set of propositions $\mathcal{A}$ is *mutually exclusive* iff the truth of any one element implies the falsehood of all other elements. The set of logical consequences of $\mathcal{A}$, written $Cn(\mathcal{A})$, is the set $\{B \subseteq W : \cap\mathcal{A} \text{ entails } B\}$. Note that if $\mathcal{A}$ is inconsistent, then $Cn(\mathcal{A})$ is $\wp(W)$, the set of all propositions over W.

A set of propositions $\mathcal{F}$ is a *field* (sometimes *algebra*) iff $\mathcal{F}$ contains W and it is closed under intersection, union and complementation. That is to say that if A, B are both elements of $\mathcal{F}$ then $W, A \cup B, A \cap B$, and $\neg A$ are also elements of $\mathcal{F}$. A set of propositions $\mathcal{F}$ is a *σ-field* (sometimes *σ-algebra*) iff it is a field that is closed under *countable* intersections, i.e. if $\mathcal{S} \subseteq \mathcal{F}$ is a countable collection of propositions, then the intersection of all its elements $\cap\mathcal{S}$ is also an element of $\mathcal{F}$. That definition implies that a σ-field is also closed under countable unions. It is not difficult to prove that an intersection of σ-fields is also a σ-field. That implies that every collection of propositions $\mathcal{F}$ generates $\sigma(\mathcal{F})$, the least σ-field containing $\mathcal{F}$, by intersecting the set of all σ-fields containing $\mathcal{F}$.

Propositions, although usually expressed by sentences in a language, are not themselves sentences. That distinction is commonly drawn by saying that propositions are *semantic* objects, whereas sentences are *syntactic* objects. Semantic objects (like propositions) are meaningful, since they represent meaningful possibilities, whereas bits of syntax must be "interpreted" before they become meaningful. In a slogan: sentences are potentially meaningful, whereas propositions already are.

For our purposes, a *language* Λ is identified with the set of all grammatical sentences it contains. Sentences will be denoted by lowercase letters $p, q, \ldots$. The language Λ is assumed to contain a set of *atomic* sentences

$a, b, \ldots$ which are not built out of any other sentences, as well as all the sentences generated by combining the atomic sentences with truth-functional connectives from propositional logic. In other words: if p, q are sentences in Λ then $\neg p$, $p \vee q$, $p \wedge q$, $p \to q$, and $p \leftrightarrow q$ are also sentences in Λ. These are meant to be read respectively as "not p," "p or q," "p and q," "if p, then q," and "p if and only if q." The symbol $\perp$ (pronounced "falsum") denotes an arbitrarily chosen contradiction (e.g. $p \wedge \neg p$) and the symbol $\top$ denotes an arbitrary tautology. Some of the sentences in Λ follow "logically" from others. For example, under the intended interpretation of the truth-functional connectives, p follows from the sentence $p \wedge q$ and also from the set of sentences $\{q, q \to p\}$. To capture the essentials of logical consequence, we introduce a *consequence operator*, which maps any set of sentences Γ to its logical consequences $Cn(\Gamma)$. The consequence operator is assumed to satisfy the following properties, which abstract the characteristic features of deductive logic.

$$\Gamma \subseteq Cn(\Gamma). \qquad \text{(Inclusion)}$$

$$\text{If } \Gamma \subseteq \Delta, \text{ then } Cn(\Gamma) \subseteq Cn(\Delta). \qquad \text{(Monotony)}$$

$$Cn(\Gamma) = Cn(Cn(\Gamma)). \qquad \text{(Idempotence)}$$

Inclusion merely expresses the triviality that any sentence p is a deductive consequence of itself. *Monotony* expresses the fact that adding more premises to a deductive argument allows you to derive all the same conclusions as you could with fewer. *Idempotence* says that $Cn(\Delta)$ contains *all* the deductive consequences of Δ. We use $\Gamma \vdash p$ as an alternative notation for $p \in Cn(\Gamma)$ and $\Gamma \nvdash p$ for $p \notin Cn(\Gamma)$. We write $\vdash p$ for $p \in Cn(\varnothing)$. The set of theorems of propositional logic is denoted by $Cn(\varnothing)$ since these can be derived from the axioms alone, without any additional assumptions.

In the following, we will sometimes assume that the consequence operator satisfies the following additional property:

$$q \in Cn(\Delta \cup \{p\}) \text{ implies } (p \to q) \in Cn(\Delta). \qquad \text{(Deduction theorem)}$$

The deduction theorem expresses the fact that you can prove the conditional sentence $p \to q$ by assuming p and then deriving q. Unsurprisingly, it is possible to prove that this property holds for most deductive logics one would encounter, including both propositional and first-order logic.

There is, of course, a systematic way to map sentences in a language to propositions. A *valuation function* V maps every atomic sentence a in Λ to a proposition $V(a)$, the set of worlds in which a is true under that interpretation of the atoms. The valuation function also interprets the non-atomic sentences in a way that respects the intended meanings of the logical connectives, i.e. so that $V(\top) = W$, $V(\neg p) = W \setminus V(p)$, and $V(p \wedge q) = V(p) \cap V(q)$. In this fashion, each sentence in Λ is mapped to a set of possible worlds. Each language Λ and valuation function V generate the field $\mathcal{F}_{\Lambda,V} = \{V(p) : p \in \Lambda\}$. In turn, $\mathcal{F}_{\Lambda,V}$ generates $\sigma(\mathcal{F}_{\Lambda,V})$, the least σ-field containing it.

We write $\Gamma \vDash p$ if for all valuations V,

$$\bigcap_{q \in \Gamma} V(q) \subseteq V(p).$$

Then, $\Gamma \vDash p$ expresses the fact that no matter how the non-logical vocabulary of Λ are interpreted, p is true in all the worlds in which all sentences in Γ are true. We say that p is *valid* iff $\{\top\} \vDash p$, i.e if $W \subseteq V(p)$ for all valuation functions. Then, p is valid iff p is true in all possible worlds, no matter how the non-logical vocabulary are interpreted. For example, the sentence $p \vee \neg p$ is valid.

We assume the following property of our deductive consequence relation.

$$\text{If } \Gamma \vdash p, \text{ then } \Gamma \vDash p. \tag{Soundness}$$

Soundness says that if the sentence p is a derivable consequence of the set of sentences Γ, then no matter how the non-logical vocabulary of Λ are interpreted, p is true in all the worlds in which all the sentences in Γ are true. That is to say that from true premises, our consequence relation always derives true conclusions. Soundness also implies that every theorem is valid. Soundness is a basic requirement of any *deductive* consequence relation, and illustrates the intended connection between deductive proof and semantic entailment.

Sentences are, in a sense, capable of expressing distinctions that propositions cannot. For example, the two sentences p and $\neg\neg p$ are obviously distinct. But if p and q are provably equivalent, i.e. if $\vdash p \leftrightarrow q$, then $\{p\} \vdash q$ and $\{q\} \vdash p$. By Soundness, $\{p\} \vDash q$ and $\{q\} \vDash p$. Therefore, for any valuation function, $V(p) = V(q)$. So p and q must express the same proposition. Of course, an agent who is unaware of the equivalence might believe p without believing q. What's worse, every sentence p such that $\vdash p$ must express the tautological proposition W. Of course, ordinary agents do not always recognize theorems of propositional logic. For this reason, some argue that it is sentences, rather than propositions, that are the appropriate objects of belief. However, most of the proposed models we will study require that rational agents adopt the same belief attitude toward logically equivalent sentences. So long as that is required, there is no significant difference between taking the objects of belief to be sentences or propositions. Still others are not satisfied with either sentences, or propositions. Perry (1979), Lewis (1979) and Stalnaker (1981) argue that in order to capture *essentially indexical* beliefs—beliefs that essentially involve indexicals such as *I*, *here*, or *now*—the objects of belief must be *centered* propositions. We will not take up this helpful suggestion here, but see Liao (2012) for a discussion of the costs and benefits of centered propositions.

2 STRUCTURES FOR FULL BELIEF

2.1 *Non-monotonic Logic*

In Section 1 we introduced the notion of a deductive consequence relation. The characteristic feature of a deductive consequence relation is that conclusions are not retracted when premises are added.

$$\text{If } \Gamma \subseteq \Delta, \text{ then } Cn(\Gamma) \subseteq Cn(\Delta). \qquad \text{(Monotony)}$$

Of course, all sorts of seemingly rational everyday reasoning violates Monotony. Reasoning according to *typicality* seems justified in ordinary circumstances, but fails to satisfy Monotony. If you were told that Tweety is a bird, you would be justified in concluding that Tweety flies, since typical birds fly. You would retract your conclusion however, if you were to learn that Tweety is a penguin. That does not mean that your original inference was unreasonable or irrational. *Inductive* inference is also famously non-monotonic. After observing one hundred white swans, you might conclude that all swans are white. Of course, you would retract your conclusion if you ever came across a black swan. *Pace* Pyrrhonian skepticism, there must be at least *some* justified inductive inferences. Ethical reasoning is also shot through with non-monotonicities. Ross (1930) discusses *prima facie* duties, or defeasible obligations, that are binding unless superseded by more urgent, and competing obligations. Ullman-Margalit (1983) points out that legal reasoning routinely relies on presumptions—of innocence, good faith, sanity, etc.—that may be withdrawn in light of new evidence. Non-Monotony is simply unavoidable in ordinary human contexts.

Non-monotonic logic studies a defeasible consequence relation $\mid\!\sim$ between premises, on the left of the wavy turnstile, and conclusions on the right. One may think of the premises on the left as a set Γ of sentences expressing "hard evidence" that an agent may possesses, and the conclusions on the right to be the defeasible conclusions that are justified on the basis of Γ. Thus, the expression $\Gamma \mid\!\sim p$ may be read as "if I were to learn all and only the sentences in Γ, I would be justified in concluding that p."

Recall from Section 1 that a deductive consequence relation satisfies Soundness, i.e. $\Gamma \vdash p$ only if p is true in all the worlds in which all sentences in Γ are true. It is clear from the preceding examples that defeasible reasoning cannot satisfy Soundness. If $\Gamma \mid\!\sim p$ then perhaps p is true in "typical" worlds in which Γ is true, or in "most" worlds in which Γ is true, or perhaps p is a sharply testable possibility compatible with Γ.

We call a consequence relation *ampliative* if $\Gamma \mid\!\sim p$, but there are worlds in which all sentences in Γ are true, but p is false. It is possible to construct consequence relations that are non-ampliative and non-monotonic, but ampliativity and non-monotonicity go hand in hand in all paradigmatic cases.

The field of artificial intelligence has, since its inception, been concerned with implementing some form of rational, ampliative, non-monotonic reasoning in artificial agents. For these purposes, deductive consequence relations are unhelpfully restrictive. That does not preclude the possibility that there is some other logic that governs good ampliative reasoning. The past forty years have seen the creation of many logics for non-monotonic inference, often developed to model a specific kind of defeasible reasoning. See Strasser and Antonelli (2018) for an excellent overview.

In view of this profusion of specialized logics, *non-monotonic logic* investigates which properties a logic of defeasible consequence must have in order to count as a *logic* at all.[1] Non-monotonic logic provides a crucial *lingua franca* for comparing different logics of defeasible inference. It is also extremely apt for the purposes of this article, because it allows us to compare different normative theories of how beliefs ought to be updated in light of new evidence, as well as theories of how full and partial beliefs ought to relate to each other.

Before we proceed to the technical development, it will be helpful to introduce an important early critique of nonmonotonic logic due to the philosopher John Pollock. Pollock (1987) identifies two sources of nonmonotonicity in defeasible reasoning. An agent may believe p, because she believes q and takes q to be a defeasible *reason* for p. Pollock distinguishes two kinds of *defeaters* for this inference: a *rebutting defeater* is a defeasible reason to believe $\neg p$, whereas an *undercutting defeater* is a reason to believe $\neg q$. Either kind of defeater may induce an agent to retract her belief in p. Pollock's point is that since nonmonotonic logics typically do not represent the structure of an agent's reasons, they often fail to elegantly handle cases of undercutting defeat. We shall soon see several examples.

2.1.1 *Principles for Nonmonotonic Logic*

Let Λ be a formal language, and let $Cn(\cdot)$ be a deductive consequence relation, as discussed in Section 1. There are in fact two closely related approaches to the study of non-montonic consequence relations. The finitary approach studies a relation between individual sentences $p \mid\!\sim q$. That approach is taken, for example, in the very influential Kraus, Lehmann, and Magidor (1990). The infinitary approach studies a relation $\Gamma \mid\!\sim p$ between an arbitrary set of sentences on the left and individual

1 Gabbay (1985) was the first to suggest this abstract point of view.

sentences on the right. That approach is taken in the canonical reference work Makinson (1994) and cannot in general be simulated by the finitary approach. For the most part we will follow Makinson (1994). However, some results are known to hold only for the finitary settings. Furthermore, the more general infinitary principles are sometimes better appreciated by their finitary consequences. For that reason, we will sometimes switch back and forth between the infinitary and the finitary approach. We write $C(\Gamma)$ for the set $\{p : \Gamma \mathrel{|\!\sim} p\}$. shorthand for $\Gamma \cup \{p\} \mathrel{|\!\sim} q$.

If defeasible logics fail to satisfy Monotony, which principles ought they satisfy? Are there some logical principles which ought to be validated by all rational defeasible reasoning? Almost all consequence relations studied in the literature satisfy the following principle.

$$\Gamma \subseteq C(\Gamma). \tag{Inclusion}$$

In its single-premise formulation Inclusion merely says that $p \mathrel{|\!\sim} p$, which is surely unexceptionable. The following two principles are also widely accepted in non-monotonic logic.

$$\Gamma \subseteq \Delta \subseteq C(\Gamma) \text{ implies } C(\Delta) \subseteq C(\Gamma). \tag{Cut}$$

$$\Gamma \subseteq \Delta \subseteq C(\Gamma) \text{ implies } C(\Gamma) \subseteq C(\Delta). \tag{Cautious Monotony}$$

As special cases, these two principles entail:

$$\Gamma \mathrel{|\!\sim} p \text{ and } \Gamma \cup \{p\} \mathrel{|\!\sim} q \text{ implies } \Gamma \mathrel{|\!\sim} q; \tag{Cut}$$

$$\Gamma \mathrel{|\!\sim} p \text{ and } \Gamma \mathrel{|\!\sim} q \text{ implies } \Gamma \cup \{p\} \mathrel{|\!\sim} q. \tag{Cautious Monotony}$$

Cut says that adding conclusions inferred from Γ to the set of premises does not *increase* inferential power. Cautious Monotony says that it does not *decrease* inferential power. If we think of the premises on the left of $\mathrel{|\!\sim}$ as my set of "hard" evidence, and the set $C(\Gamma)$ as a theory inductively inferred on the basis of Γ, then Cautious Monotony is an expression of hypothetico-deductivism: if I observe a consequence of my theory $C(\Gamma)$, I should not thereby retract any previous conclusions. Moreover, Cut says that I should not add any new conclusions. Taken together the two principles say that if you observe a consequence of your theory, you should not change it:

$$\Gamma \subseteq \Delta \subseteq C(\Gamma) \text{ implies } C(\Gamma) = C(\Delta). \tag{Cumulativity}$$

Gabbay (1985) proposes that (finitary versions of) Inclusion, Cut and Cautious Monotony are the minimal properties that every interesting non-monotonic logic must satisfy. That remains the consensus view to this day. It is easy to show that Inclusion and Cut jointly imply a principle familiar from Section 1:

$$C(\Gamma) = C(C(\Gamma)). \qquad \text{(Idempotence)}$$

There is also the question of how a non-monotonic consequence relation $C(\cdot)$ should interact with a classical relation of deductive consequence $Cn(\cdot)$. The following principle says that defeasible reasoning allows you to make strictly more conclusions than classical deductive reasoning:

$$Cn(\Gamma) \subseteq C(\Gamma). \qquad \text{(Supraclassicality)}$$

That is perhaps unreasonable if we think of $C(\cdot)$ as modeling the defeasible reasoning of some bounded agent. It begins to sound better if we think of $C(\Gamma)$ as modeling the ampliative conclusions that are justified on the basis of Γ.

Makinson (1994) observes that any supraclassical $C(\cdot)$ that satisfies Idempotence and Cumulativity also satisfies the following pair of principles.

$$Cn(C(\Gamma)) = C(\Gamma). \qquad \text{(Left Absorption)}$$

$$C(\Gamma) = C(Cn(\Gamma)). \qquad \text{(Right Absorption)}$$

Left Absorption says that $C(\Gamma)$ is closed under deductive consequence. Right Absorption says that the conclusions that are justified on the basis of Γ depend only on the logical content of Γ, and not on its mode of presentation. The conjunction of Right and Left Absorption is called Full Absorption.

Makinson advocates for one more interaction principle:

$$C(\Gamma) \cap C(\Delta) \subseteq C(Cn(\Gamma) \cap Cn(\Delta)). \qquad \text{(Distribution)}$$

That condition is perhaps too complex to admit of an intuitive gloss. However, we can better understand its meaning from its finitary consequences. Any supraclassical consequence relation satisfying Distribution and Full Absorption also satisfies the following.

$$\Gamma \cup \{p\} \mid\!\sim r \text{ and } \Gamma \cup \{q\} \mid\!\sim r \text{ implies } \Gamma \cup \{p \vee q\} \mid\!\sim r. \qquad \text{(Or)}$$

$$\Gamma \cup \{p\} \mid\!\sim q \text{ and } \Gamma \cup \{\neg p\} \mid\!\sim q \text{ implies } \Gamma \mid\!\sim q. \qquad \text{(Case reasoning)}$$

These two principles seem to be very compelling. Any genuine consequence relation ought to enable reasoning by cases. If I would infer q irrespective of what I learned about p, I should be able to infer q before the matter of p has been decided. Similarly, if p follows defeasibly from both p and q, it ought to follow from their disjunction. Any consequence relation that satisfies Supraclassicality, Left Absorption and Case Reasoning must also satisfy the following principle:

$$\Gamma \cup \{p\} \mid\!\sim q \text{ implies } \Gamma \mid\!\sim p \to q. \qquad \text{(Conditionalization)}$$

To prove that entailment suppose that $\Gamma \cup \{p\} \mid\sim q$. Since $p \to q$ is a deductive consequence of q, it follows by Left Absorption that $\Gamma \cup \{p\} \mid\sim p \to q$. Furthermore, since $p \to q$ is a deductive consequence of $\neg p$ it follows by supraclassicality that $\Gamma \cup \{\neg p\} \mid\sim p \to q$. By Case Reasoning, $\Gamma \mid\sim p \to q$.

Conditionalization says that upon learning new evidence, you never "jump to conclusions" that are not entailed by the deductive closure of your old beliefs with the new evidence. That is not an obviously appealing principle. An agent that starts out with $\Gamma = Cn(\varnothing)$ will either fail to validate Conditionalization or never make any ampliative inferences at all. Suppose that after observing 100 black ravens an agent validating Conditionalization comes to believe that all ravens are black. Then, at the outset of inquiry, she must have believed that either all ravens are black, or she will see the first non-black raven among the first hundred. Such an agent seems strangely opinionated about when the first counterexample to the inductive generalization must appear.

For a more realistic example, consider the 1887 Michelson-Morely experiment. After a null result failing to detect any *significant* difference between the speed of light in the prevailing direction of the presumed aether wind, and the speed at right angles to the wind, physicists turned against the aether theory. If the physicists validated Conditionalization then, before the experiments, they must have believed that either there is no luminiferous aether, or the aether wind blows quickly enough to be detected by their equipment. But why should they have been so confident that the aether wind is not too slow to be detectable? Even if there is nothing objectionable about an agent who validates Conditionalization, there is something very anti-inductivist about the thesis that *all* justified defeasible inferences on the basis of new evidence can be reconstructed as deductive inferences from prior conclusions plus the new evidence. Schurz (2011) makes a similar criticism, in a slightly different context:

> Inductive generalizations as well as abductive conjectures accompany belief expansions by new observations, in science as well as in common sense cognitions. After observing several instances of a 'constant conjunction,' humans almost automatically form the corresponding inductive generalization; and after performing a new experimental result sufficiently many times, experimental scientists proclaim the discovery of a new empirical law ... [Conditioning]-type expansion is not at all creative but merely additive: it simply adds the new information and forms the deductive closure, but never generates new (non-logically entailed) hypotheses.

Schurz objects that, according to Conditionalization, dispositions to form inductive generalizations must be "programmed in" with material conditionals at the outset of inquiry. Anyone sympathetic to this view must reject either Supraclassicality, Left Absorption, or Case Reasoning. Finding such surprising consequences of seemingly unproblematic principles is one of the boons of studying non-monotonic logic.

We finish this section by introducing one more prominent and controversial principles of non-monotonic logic. The position one takes on this principle will determine how one feels about many of the theories which we turn to in the following. Kraus et al. (1990) claim that any rational reasoner should validate the following strengthening of Cautious Monotony.

$$\Gamma \mid\!\sim p \text{ and } \Gamma \mid\!\not\sim \neg q \text{ entails } \Gamma \cup \{q\} \mid\!\sim p. \qquad \text{(Rational Monotony)}$$

Rational Monotony says that so long as new evidence q is logically compatible with your prior beliefs $C(\Gamma)$, you should not retract any beliefs from $C(\Gamma)$. Accepting both Rational Monotony and Conditionalization amounts to saying that when confronted with new evidence that is logically consistent with her beliefs, a rational agent responds by simply forming the deductive closure of her existing beliefs with the new evidence. On that view, deductive logic is the only necessary guide to reasoning, so long as you do not run into contradiction. Stalnaker (1994) gives the following well-known purported counterexample to Rational Monotony.

Suppose an agent initially believes the following about the three composers Verdi, Bizet, and Satie.

(Iv) Verdi is Italian;

(Fb) Bizet is French;

(Fs) Satie is French.

Let p be the sentence that Verdi and Bizet are compatriots, i.e. $(Fv \wedge Fb) \vee (Iv \wedge Ib)$. Let q be the sentence that Bizet and Satie are compatriots. Suppose that the agent receives the evidence p. As a result, she retracts her belief in $Iv \wedge Fb$ concluding that either Verdi and Bizet are both French or they are both Italian. She retains her belief that Satie is French. Notice that after updating on p, she believes it is possible that Bizet and Satie are compatriots, i.e. $p \mid\!\not\sim \neg q$. Now suppose that she receives the evidence q. Since q is compatible with all her previous conclusions, Rational Monotony requires her to conclude that all three composers are French. However, it seems perfectly rational to suspend judgment and concludes that the three are either all Italian, or all French.

Kelly and Lin (forthcoming) give the following counterexample to Rational Monotony, based on Lehrer's (1965) no-false-lemma variant of Gettier's famous (1963) scenario. There are just two people in your office, named Alice and Bob. You are interested in whether one of them owns a certain Ford. Let p be the sentence that Alice owns the Ford. Let q be the sentence that Bob has the Ford. You have inconclusive evidence that Alice owns the Ford—you saw her driving one just like it. You have weaker evidence that Bob owns the Ford—his brother owns a Ford dealership. Based on that evidence Γ you conclude $p \vee q$, i.e. that someone in the office owns the Ford, but do not go so far as inferring p, or q. You ask Alice and she tells you that the Ford she was driving was rented. That defeats your main *reason* for $p \vee q$, therefore you retract your belief that someone in the office has a Ford. But since $\Gamma \not\mid\!\sim \neg p$, Rational Monotony requires you to conclude that Bob owns the Ford. However, there does not seem to be anything irrational about how you have reasoned. This seems to be an illustration of Pollock's (1987) point: the logic is going wrong because it is ignoring the structure of the agent's reasons.

We end this section on a terminological note. It is common in the literature to use *System P* (Preferential) to refer to the following set of single-premise principles, labeled so that the reader can identify their infinitary analogues. The terminology is due to Kraus et al. (1990).

$$p \mid\!\sim p. \tag{Reflexivity}$$

$$\vdash p \leftrightarrow q \text{ and } p \mid\!\sim r \text{ implies } q \mid\!\sim r. \tag{Left equivalence}$$

$$\vdash q \rightarrow r \text{ and } p \mid\!\sim q \text{ implies } p \mid\!\sim r. \tag{Right weakening}$$

$$p \mid\!\sim q \text{ and } p \mid\!\sim r \text{ implies } p \mid\!\sim q \wedge r. \tag{And}$$

$$p \mid\!\sim r \text{ and } q \mid\!\sim r \text{ implies } p \vee q \mid\!\sim r. \tag{Or}$$

$$p \mid\!\sim q \text{ and } p \mid\!\sim r \text{ implies } p \wedge q \mid\!\sim r. \tag{Cautious monotony}$$

System R (Rational) arises from System P by adding a single-premise version of Rational Monotony:

$$p \mid\!\sim r \text{ and } p \not\mid\!\sim \neg q \text{ implies } p \wedge q \mid\!\sim r. \tag{Rational monotony}$$

2.1.2 *Preferential Semantics*

So far we have considered a non-monotonic consequence relation merely as a relation between syntactic objects. We can rephrase properties of non-monotonic logic "semantically," i.e. in terms of the possible worlds in which the sentences are true or false. In some cases, this allows us to give a very perspicuous view on defeasible logic.

Recall from Section 1 that a deductive consequence relation satisfies Soundness, i.e. that $\Gamma \vdash p$ only if p is true in all the worlds in which all sentences in Γ are true. As we have discussed, non-monotonic logics are ampliative, and therefore must violate Soundness. Shoham (1987) inaugurated a semantics for non-monotonic logics in which $\Gamma \mid\!\sim p$ only if p is true in a "preferred" set of worlds in which Γ is true. These are usually interpreted as the "most typical," or "most normal" worlds in which all sentences in Γ are true. If Γ is a set of sentences in Λ and V is a valuation function, we write $V(\Gamma)$ as shorthand for $\cap_{q \in \Gamma} V(q)$. See Section 1 if you need a refresher on valuation functions. Kraus et al. (1990) first proved most of the results of this section for single-premise consequence relations. We follow Makinson (1994) in presenting their infinitary generalizations.

A *preferential model* is a triple $\langle W, V, < \rangle$ where W is a set of possible worlds, V is a valuation function and $<$ is an arbitrary relation on the elements of W. The relation $<$ is *transitive* iff $x < y$ and $y < z$ implies $x < z$. The relation $<$ is *irreflexive* iff for all $w \in W$ it is not the case that $w < w$. A transitive, irreflexive relation is called a *strict order*. We write $w \leq v$ iff $w < v$ or $w = v$. The strict order $<$ is *total* iff for $w, v \in W$ either $w \leq v$ or $v \leq w$.

If Γ is a set of sentences, we say that $w \in \text{Min}_<(\Gamma)$ iff $w \in V(\Gamma)$ and there is no $v \in V(\Gamma)$ such that $v < w$. In other words, $w \in \text{Min}_<(\Gamma)$ iff w is a $<$-minimal element of $V(\Gamma)$. Every preferential model gives rise to a consequence relation by letting

$$\Gamma \mid\!\sim_< p \text{ iff } \text{Min}_<(\Gamma) \subseteq V(p),$$

i.e. $\Gamma \mid\!\sim_< p$ iff p is true in all the *minimal* worlds in which all sentences in Γ are true. Write $C_<(\Gamma)$ for the set $\{p : \Gamma \mid\!\sim_< p\}$.

We say that a preferential model is *stoppered* iff for every set of sentences Γ, if $w \in V(\Gamma)$ then there is $v \leq w$ such that $v \in \text{Min}_<(\Gamma)$. (Note that Kraus et al., 1990, called stoppered models *smooth* models.) Makinson (1994) proves the following.

Theorem 1 *Suppose that $\mathfrak{M} = \langle W, V, < \rangle$ is a preferential model. Then $C_<(\cdot)$ satisfies Inclusion, Cut, Supraclassicality, and Distribution. If $\mathfrak{M}$ is stoppered, then $C_<(\cdot)$ also satisfies Cautious Monotony.*

Makinson (1994) also gives the following two partial converses. The latter essentially reports a result from Kraus et al. (1990).

Theorem 2 *If $C(\cdot)$ satisfies Inclusion, Cut, and Cautious Monotony, there is a stoppered preferential model $\mathfrak{M} = \langle W, V, < \rangle$ such that $C(\cdot) = C_<(\cdot)$.*

Theorem 3 *If $C(\cdot)$ satisfies Inclusion, Cut, Cautious Monotony, Supraclassicality, and Distribution, then there is a stoppered preferential model $\mathfrak{M} = \langle W, V, < \rangle$ such that for all finite $\Delta \subseteq \Lambda$, $C(\Delta) = C_<(\Delta)$. Moreover, $\mathfrak{M}$ may be constructed such that $<$ is a strict order.*

Taken together, Theorem 1 and Theorem 3 say that, at least for finitary consequences, the consequence relations generated by preferential models are exactly the consequence relations satisfying Inclusion, Cut, Cautious Monotonicity, Supraclassicality, and Distribution. In fact, one can always think of these preferential models as generated by a strict (partial) order. The question remains whether there are any natural conditions on preferential models that ensure that Rational Monotony is also satisfied. It turns out that Rational Monotony follows from the requirement that the preference relation $<$ is a total order.

Say that a preferential model $\mathfrak{M} = \langle W, V, < \rangle$ is *modular* iff for all $w, u, v \in W$, if $w < u \approx v$ then $w < v$. Here $u \approx v$ means that u, v are unordered, i.e. it is not the case that $u < v$ and it is not the case that $v < u$. If $<$ is a strict order, modularity is equivalent to the intuitive property of *rankedness*: there is a totally ordered set T and a function $\rho : W \to T$ such that for all $u, v \in W$, $u < v$ iff $\rho(u) \ll \rho(v)$, where $\ll$ is the total ordering of T. Makinson proves the following.

Theorem 4 *Suppose that $\mathfrak{M} = \langle W, V, < \rangle$ is a preferential model. If $\mathfrak{M}$ is modular, then $C_<(\cdot)$ satisfies Rational Monotony.*

Kraus et al. (1990) prove the following partial converse.

Theorem 5 *If $C(\cdot)$* finitarily *satisfies Inclusion, Cut, Cautious Monotony, Supraclassicality, Distribution, and Rational Monotony, then there is a ranked, stoppered preferential model $\mathfrak{M} = \langle W, V, < \rangle$ such that for all finite $\Delta \subseteq \Lambda$, $C(\Delta) = C_<(\Delta)$. Moreover, $\mathfrak{M}$ may be constructed such that $<$ is a strict order.*

The essential difference between preferential models that satisfy Rational Monotony and those that do not is that the former correspond to those generated by a *ranked* partial order. This result is helpful to keep in mind because in the following we will see several models of belief that can be understood as arising from a *total* plausibility order, and some that arise from a merely *partial* plausibility order. In light of Theorem 3 and Theorem 5, we can expect the former to satisfy System R and the latter to satisfy only the weaker System P.

2.2 *AGM Belief Revision Theory*

The theory of belief revision is concerned with how to update one's beliefs in light of new evidence, especially when new evidence is inconsistent with prior beliefs. It is especially occupied with the following sort of scenario, borrowed from Gärdenfors (1992). Suppose that you believe all the following sentences:

(*a*) All European swans are white;

(b) The bird in the pond is a swan;

(c) The bird in the pond comes from Sweden;

(d) Sweden is in Europe.

Now suppose that you were to learn the sentence e that the bird in the pond is black. Clearly, e is inconsistent with your beliefs a, b, c, d. If you want to incorporate the new information e and remain consistent, you will have to retract some of your original beliefs. The problem of belief revision is that deductive logic alone cannot tell you which of your beliefs to give up—this has to be decided by some other means. Considering a similar problem, Quine and Ullian (1970) enunciated the principle of "conservatism," counseling that our new beliefs "may have to conflict with some of our previous beliefs; but the fewer the better." In his (1990), Quine dubs this the "maxim of minimal mutilation." Inspired by these suggestive principles, Alchourrón, Gärdenfors, and Makinson (1985) develop a highly influential theory of belief revision, known thereafter as AGM theory, after its three originators.

In AGM theory, beliefs held by an agent are represented by a set B of sentences. The set B is called the *belief state* of the agent. This set is usually assumed to be closed under logical consequence. Of course, this is an unrealistic idealization, since it means that the agent believes all logical consequences of her beliefs. Levi (1991) defends this idealization by changing the interpretation of the set B—these are the sentences that the agent is *committed* to believe, not those that she actually believes. Although we may never live up to our commitments, Levi argues that we are committed to the logical consequences of our beliefs. That may rescue the principle, but only by changing the interpretation of the theory.

AGM theory studies three different types of belief change. *Contraction* occurs when the belief state B is replaced by $B \div p$, a logically closed subset of B no longer containing p. *Expansion* occurs when the belief state B is replaced with $B + p = Cn(B \cup \{p\})$, the result of simply adding p to the set of beliefs and closing under logical consequence. *Revision* occurs when the belief state B is replaced by $B * p$, the result of adding p to B and removing whatever is necessary to ensure that the resulting belief state $B * p$ is logically consistent.

Contraction is the fundamental form of belief change studied by AGM. There is no mystery in how to define expansion, and revision is usually defined derivatively via the *Levi identity* (1977): $B * p = (B \div \neg p) + p$. Alchourrón et al. (1985) and Gärdenfors and Makinson (1988) proceed axiomatically: they postulate several principles that every rational contraction operation must satisfy. Fundamental to AGM theory are several representation theorems showing that certain intuitive constructions give rise to contraction operations satisfying the basic postulates and conversely, that every operation satisfying the basic postulates can be seen as the outcome of such a construction. See Lin (this volume) for an introduction to these results.

AGM theory is unique in focusing on belief contraction. For someone concerned with maintaining a database, contraction is a fairly natural operation. Medical researchers might want to publish a data set, but make sure that it cannot be used to identify their patients. Privacy regulations may force data collectors to "forget" certain facts about you and, naturally, they would want to do this as conservatively as possible. However, a plausible argument holds that all forms of rational belief change occurring "in the wild" involve learning new information, rather than conservatively removing an old belief. All the other formalisms covered in the article focus on this form of belief change. For this reason, we focus on the AGM theory of revision and neglect contraction.

Before delving into some of the technical development, we mention some important objections and alternatives to the AGM framework. As we have mentioned, the belief state of an agent is represented by the (deductively closed) set B of sentences the agent is committed to believe. The structure of the agent's *reasons* is not represented: you cannot tell of any two $p, q \in B$ whether one is a reason for the other. Gärdenfors (1992) distinguishes between *foundations* theories, that keep track of which beliefs justify which others, and *coherence* theories, which ignore the structure of justification and focus instead on whether beliefs are consistent with one another. Arguing for the coherence approach, Gärdenfors (1992) draws a stark distinction between the two:

> According to the foundations theory, belief revision should consist, first, in giving up all beliefs that no longer have a *satisfactory justification* and, second, in adding new beliefs that have become justified. On the other hand, according to the coherence theory, the objectives are, first, to maintain *consistency* in the revised epistemic state, and, second, to make *minimal changes* of the old state that guarantee overall coherence.

Implicit in this passage is the idea that foundations theory are fundamentally out of sympathy with the principle of minimal mutilation. Elsewhere (1988), Gärdenfors is more apologetic, suggesting that some hybrid theory is possible and perhaps even preferable:

> I admit that the postulates for contractions and revisions that have been introduced here are quite simpleminded, but they seem to capture what can be formulated for the meager structure of belief sets. In richer models of epistemic states, admitting, for example, reasons to be formulated, the corresponding

> conservativity postulates must be formulated much more cautiously (p. 67).

Previously, we have seen Pollock (1987) advocating for foundationalism. In artificial intelligence, Doyle's (1979) *reason maintenance system* is taken to exemplify the foundations approach. Horty (2012) argues that default logic aptly represents the structure of reasons. For a defense of foundationalism, as well as a useful comparison of the two approaches, see Doyle (1992).

Another dissenting tradition advocates for *belief bases* instead of belief states. A belief base is a set of sentences that is typically not closed under logical consequence. Its elements represent "basic" beliefs that are not derived from other beliefs. This allows us to distinguish between sentences that are explicit beliefs, like "Shakespeare wrote Hamlet" and never thought-of consequences like "Either Shakespeare wrote Hamlet or Alan Turing was born on a Monday." Revision and contraction are then redefined to operate on belief bases, rather than belief sets. That allows for increased expressive power, since belief bases which have the same logical closure are not treated interchangeably. For an introduction to belief bases see Hansson (2017). For a book-length treatment, see Hansson (1999).

Finally, one of the most common criticisms of AGM theory is that it does not illuminate *iterated* belief change. In the following, we shall see that the canonical revision operation takes as input an entrenchment ordering on a belief state, but outputs a belief state without an entrenchment order. That severely underdetermines the result of a subsequent revision. For more on the problem of iterated belief revision, see Huber (2013a).

The treatment in this article is necessarily rather compressed. There are several excellent survey articles on belief revision. See Hansson (2017), Huber (2013a, 2013b), and Lin (this volume).

2.2.1 *Revision*

Alchourrón et al. (1985) propose the following postulates for rational belief revision.

$$B * p = Cn(B * p). \tag{Closure}$$

$$p \in B * p. \tag{Success}$$

$$B * p \subseteq Cn(B \cup \{p\}). \tag{Inclusion}$$

$$\text{If } \neg p \notin Cn(B), \text{ then } B \subseteq B * p. \tag{Preservation}$$

$$B * p \text{ is consistent if } p \text{ is consistent.} \tag{Consistency}$$

$$\text{If } (p \leftrightarrow q) \in Cn(\varnothing), \text{ then } B * p = B * q. \tag{Extensionality}$$

By now, Closure, Success, Consistency, and Extensionality should be straightforward to interpret. These postulates impose synchronic constraints on $B * p$. Preservation and Inclusion are the only norms that are really about *revision*—they capture the diachronic spirit of AGM revision. Inclusion says that revision by p should yield *no more* new beliefs than expansion by p. In other words, any sentence q that you come to believe after revising by p is a deductive consequence of p and your prior beliefs. Consider the following principle:

$$\text{If } q \in B * p, \text{ then } (p \to q) \in B. \qquad \text{(Conditionalization)}$$

In Section 2.1.1, we considered an analogue of Conditionalization for nonmonotonic logic. All the same objections apply equally well in the context of belief revision. Recall from Section 1 that a deductive consequence relation admits a deduction theorem iff $\Delta \cup \{p\} \vdash q$ implies that $\Delta \vdash p \to q$. So long as a deduction theorem is provable for $Cn(\cdot)$, Inclusion and Conditionalization are equivalent. To see this, suppose that the revision operation $*$ satisfies Inclusion. Then, if $q \in B * p$, it follows that $B \cup \{p\} \vdash q$. By the deduction theorem, $B \vdash p \to q$. For the converse, suppose that the revision operation $*$ satisfies Conditionalization. Then, if $q \in B * p$, it follows that $p \to q \in B$ and $q \in Cn(B \cup \{p\})$. If you found any of the arguments against Conditionalization convincing, you ought to be skeptical of Inclusion.

Preservation says that, so long as the new information p is logically consistent with your prior beliefs, all of your prior beliefs survive revision by p. In the setting of non-monotonic logic, we called this principle Rational Monotony. All objections and counterexamples to Rational Monotony from Section 2.1.1 apply equally well in belief revision. As we have seen, Preservation rules out any kind of *undercutting* defeat of previously successful defeasible inferences. Accepting both Preservation (Rational Monotonicity) and Inclusion (Conditionalization) amounts to saying that when confronted with new evidence that is logically consistent with her beliefs, a rational agent responds by simply forming the deductive closure of her existing beliefs with the new evidence. On that view, deductive logic is the only necessary guide to reasoning, so long as you do not run into contradiction.

Alchourrón et al. (1985) also propose the following supplementary revision postulates, closely related to Inclusion and Preservation.

$$B * (p \wedge q) \subseteq (B * p) + q. \qquad \text{(Conjunctive Inclusion)}$$

$$\text{If } \neg q \notin Cn(B * p), \text{ then } (B * p) + q \subseteq B * (p \wedge q). \qquad \text{(Conjunctive Preservation)}$$

It is possible to make the connection between belief revision and nonmonotonic logic precise. Given a belief set B and a revision operation $*$, we can define a single-premise consequence relation by setting

$$p \mid\!\sim q \text{ iff } q \in B * p.$$

Similarly, given a single-premise consequence relation $\mid\!\sim$ we can define

$$B = \{p : \top \mid\!\sim p\} \text{ and } B * p = \{q : p \mid\!\sim q\}.$$

Then it is possible to prove the following correspondences between AGM belief revision and the set of single-premise principles we called System R in Section 2.1.1. It follows, by Theorem 5, that AGM revision can be represented in terms of a ranked, stoppered preferential model over possible worlds.

Theorem 6 *Suppose that $*$ is a revision operation for B satisfying all eight revision postulates. Then, the nonmonotonic consequence relation given by $p \mid\!\sim q$ iff $q \in B * p$ satisfies all the principles of System R.*

Theorem 7 *Suppose that $\mid\!\sim$ is a consequence relation that satisfies all the principles of System R and such that $p \mid\!\sim \bot$ only if $\vdash \neg p$. Then, the revision operation $*$ defined by letting $B = \{p : \top \mid\!\sim p\}$ and $B * p = \{q : p \mid\!\sim q\}$ satisfies all eight revision postulates.*

2.2.2 *Entrenchment*

Gärdenfors and Makinson (1988) introduce the notion of an *entrenchment relation* on sentences.

> Even if all sentences in a [...] set are accepted or considered as facts [...], this does not mean that all sentences are of equal value for planning or problem-solving purposes. Certain [...] beliefs about the world are more important than others when planning future actions, conducting scientific investigations, or reasoning in general. We will say that some sentences [...] have a higher degree of *epistemic entrenchment* than others. The degree of entrenchment will, intuitively, have a bearing on what is abandoned [...], and what is retained, when a contraction or revision is carried out.

To model the degree of entrenchment, Gärdenfors and Makinson (1988) introduce a relation $\leq$ holding between sentences of the language Λ. The notation $p \leq q$ is pronounced "p is at most as entrenched as q." Gärdenfors and Makinson (1988) propose that the entrenchment relation $\leq$ satisfy the following postulates.

If $p \leq q$ and $q \leq r$, then $p \leq r$. (Transitivity)

If $p \vdash q$, then $p \leq q$. (Dominance)

Either $p \leq (p \wedge q)$, or $q \leq (p \wedge q)$. (Conjunctiveness)

If B is consistent, then $p \notin B$ iff $p \leq q$ for all q. (Minimality)

If $q \leq p$ for all q, then $p \in Cn(\varnothing)$. (Maximality)

Note that, in light of Minimality, an entrenchment relation is defined for a particular belief set B. It follows from the first three of these postulates that an entrenchment order is *total*, i.e. for all p, q either $p \leq q$ or $q \leq p$.

Given a belief set B and an entrenchment relation $\leq$, it is possible to define a revision operation directly by setting:

$$B * p = Cn(\{q \in \Lambda : \neg p < q\} \cup \{p\}). \qquad (C*)$$

The idea behind this equation is that the agent revises by p by first clearing from her belief set anything less entrenched than $\neg p$, (by dominance, this includes everything entailing $\neg p$) adding p, and then closing under logical consequence. This illustrates why AGM theory is not a theory of *iterated* revision: the revision operation takes as input an entrenchment order and belief state, but outputs only a belief state. That severely underdetermines the results of subsequent revisions. Gärdenfors (1988) proves the following.

Theorem 8 *If a relation $\leq$ satisfies the five entrenchment postulates, then the revision function $*$ determined via $(C*)$ satisfies the six basic and the two supplementary revision postulates.*

Finally, given a belief set B, an entrenchment relation can be recovered from a revision operation by setting:

$$p \leq q \text{ iff } p \notin B * \neg(p \wedge q) \text{ or } \vdash q. \qquad (C^*_{\leq})$$

The idea is that p is no more entrenched than q if p does not survive a revision by $\neg(p \wedge q)$ or if q is a tautology. Rott (2003) proves the following.

Theorem 9 *If a revision operation $*$ satisfies the six basic and the two supplementary contraction postulates, then the entrenchment relation determined via $(C^*_{\leq})$ satisfies the five entrenchment postulates.*

2.2.3 *Sphere Semantics*

So far we have thought of belief revision syntactically: a revision operation $*$ takes in a set B of syntactic objects and a sentence p and outputs another set of sentences $B * p$. Grove (1988) gives a perspicuous way to represent the revision postulates semantically, i.e. in terms of the possible worlds in which the sentences are true or false.

As before, let W be a set of possible worlds and let $V : \Lambda \to \wp(W)$ be a valuation function.[2] If Γ is a set of sentences in Λ, we write $V(\Gamma)$ as shorthand for $\cap_{q \in \Gamma} V(q)$. If E is a proposition, we write $T(E)$ as shorthand for $\{p \in \Lambda : E \subseteq V(p)\}$, i.e. the set of all sentences p such that E entails $V(p)$.

A set of propositions $\mathcal{S}$ is a *system of spheres* centered on $V(B) \subseteq W$ iff for all $E, F \subseteq W$ and all $p \in \Lambda$, the following conditions hold.

$$\text{If } E, F \in \mathcal{S}, \text{ then } E \subseteq F \text{ or } F \subseteq E. \tag{Nested}$$

$$V(B) \in \mathcal{S} \text{ and if } E \in \mathcal{S} \text{ then } V(B) \subseteq E. \tag{Centered}$$

$$W \in \mathcal{S}. \tag{Maximum}$$

$$\begin{aligned}&\text{If } p \nvdash \bot, \text{ then there is } E \in \mathcal{S} \text{ such that}\\ &\qquad E \cap V(p) \neq \varnothing \text{ and if } F \cap V(p) \neq \varnothing \text{ then } E \subseteq F.\end{aligned} \tag{Well order}$$

In other words, a system of spheres centered on $V(B)$ is a nested set of propositions, all entailed by $V(B)$, with the following property: if p is a consistent sentence, then there is a logically strongest element of $\mathcal{S}$ consistent with $V(p)$. If p is a consistent sentence, let $\mathcal{S}(p)$ be $E \cap V(p)$, where E is the logically strongest element of $\mathcal{S}$ consistent with $V(p)$. Otherwise, let $\mathcal{S}(p) = \varnothing$. In other words: $\mathcal{S}(p)$ is the set of worlds compatible with $V(p)$ that is "closest" to $V(B)$ according to the sphere system. Note that if $V(p) \cap V(B) \neq \varnothing$, then $\mathcal{S}(p) = V(B) \cap V(p)$. If $V(p) \cap V(B) = \varnothing$ we find the closest sphere compatible with $V(p)$ and intersect the two. Given a belief set B and a system of spheres $\mathcal{S}$ centered on $V(B)$ we can define a revision operator by setting:

$$B * p = T(\mathcal{S}(p)).$$

The idea is this: when you revise on a sentence p compatible with your previous beliefs, then the strongest proposition you believe is $V(p) \cap V(B)$. If p is incompatible with your beliefs, you fall back to $E \cap V(p)$, where E is the p-compatible proposition closest to your old belief $V(B)$. Thus, the system of spheres $\mathcal{S}$ can be seen as a set of "fallback positions" for updating on incompatible propositions. See Figure 1.

Grove (1988) proves the following.

Theorem 10 *Let B be a belief set. For each system of spheres $\mathcal{S}$ centered on $V(B)$, there is an operation $*$ satisfying the six basic and the two supplementary revision postulates such that $B * p = T(\mathcal{S}(p))$. Moreover, for every revision operation $*$ satisfying the six basic and the two supplementary revision postulates, there is a sphere system $\mathcal{S}$ centered on $V(B)$ such that $B * p = T(\mathcal{S}(p))$.*

2 See Section 1 if you need a refresher on valuation functions.

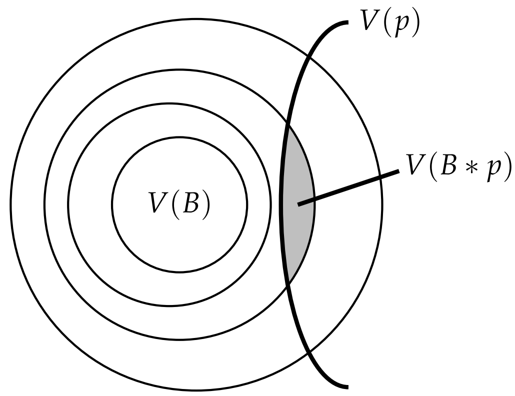

Figure 1: A system of spheres centered on $V(B)$. The shaded region is $V(B * p)$.

Finally, the sphere semantics gives a dramatic illustration of the critique that AGM does not illuminate iterated belief change. A revision maps a set of spheres $\mathcal{S}$ and a sentence p to a *set of sentences* $T(\mathcal{S}(p))$. It does not output a new set of spheres centered on the new belief state. That severely underdetermines the result of future revisions.

2.3 *The Paradox of the Preface*

The models of full belief that we have seen so far require that beliefs be consistent and closed under deductive consequence. While it is admitted that this requirement is not psychologically realistic, or perhaps even feasible for bounded agents, it is proffered as a normative principle that we should strive to approximate. After all, consistency and closure are *necessary* conditions for achieving the following two related ends: believing only true sentences (consistency) and believing as many true sentences as possible without risking error in any more possible worlds (closure).

Nevertheless, the Paradox of the Preface, due to Makinson (1965), challenges even the *normativity* of deductive consistency. The story goes like this. A famous theorist has just finished her latest book. As is customary for such works, she includes a passage in the preface thanking her colleagues and students for their help in editing and proofreading, but accepting sole responsibility for the mistakes that inevitably remain. She seems to be saying that, despite her best efforts, she believes that not everything that she asserts in the book is true. Let $s_1, \ldots, s_n$ be the claims she asserts in the book. Presumably, she believes each of the s_i or else she would not have asserted them. Yet in the preface she claims to believe $\neg(s_1 \wedge \ldots \wedge s_n)$, the claim that at least one of the s_i is false. The theorist seems to be behaving perfectly rationally, yet on this reconstruction there is no way that she can be both consistent and deductively closed.

It is tempting to say that inconsistency in the service of intellectual humility is no sin. Yet this creates a further difficulty: surely some inconsistencies are vicious and should be eliminated. Others, it seems, are virtuous and ought to be tolerated. But how do we know which inconsistencies are which? If we point out an inconsistency in someone's beliefs, we tend to think that they are under some pressure to resolve it. But why can't everyone respond to such a challenge by claiming that their inconsistency is a virtue? The Preface Paradox seems to challenge the very normativity of logical consistency.

There are several ways to respond to this challenge. The first, and perhaps the most common route, is to claim that belief is fundamentally a matter of degree. The theorist merely has a high *degree* of belief in each of the statements of her book. And there is nothing surprising about having a high degree of belief in each of the s_i but not in their conjunction. In fact, if the structure of partial belief is probabilistic, this would emerge as a simple consequence of the probability calculus: it is to be expected that the probability of each of the s_i exceeds the probability of their conjunction, so long as the probability of the s_i falls short of unity. This analysis also entails something about the relationship between full and partial belief: it is rationally admissible to fully believe statements that have a high, but not maximal, degree of belief. These themes will be taken up in subsequent sections.

A second set of responses to the paradox calls our attention to the variety of cognitive attitudes that are involved in the story. For example, Cohen (1992) attributes many confusions and apparent paradoxes to the erroneous conflation of two related cognitive attitudes: belief and *acceptance*. Belief in p, according to Cohen, is a disposition to feel it true that p, whenever attending to issues raised by the proposition that p. This disposition may or may not be reflected in speech and action, and is not under direct volitional control. But to *accept* that p is to adopt a policy of deeming, positing, or postulating that p—i.e. of including it among one's premises for deciding what to do or think in a particular context, whether or not one feels it to be true. Acceptance is a volitional matter, and is sensitive to our cognitive context.

Belief is sometimes not even a *prima facie* reason for acceptance: in scientific contexts many of our most cherished beliefs are not accepted, as it is our duty to subject them to criticism and not argue from them as premises. Cohen claims that a person who accepts nothing that she believes is intellectually paralyzed, but someone who accepts everything she believes is recklessly uncritical. Furthermore, acceptance may sometimes promote belief, at least in the long run, but often has no effect: for example, a defense lawyer may accept that her client is innocent, but believe otherwise.

Cohen claims that acceptance ought to be closed, at least under accepted deductive consequences. Consistency is also, presumably, a norm of acceptance. Belief, however, is different:

> you are not intellectually pledged by a set of *beliefs*, however strong, to each deductive consequence of that set of beliefs, even if you recognize it to be such. That is because belief that p is a disposition to feel that p, and feelings that arise in you [...] through involuntary processes [...] no more impose their logical consequences on you than do the electoral campaign posters that people stick on your walls without your consent. (Cohen, 1992, p. 31)

Armed with the distinction between belief and acceptance, we can attempt a redescription of the preface paradox. In the context of her theoretical work, the theorist accepts $s_1, \ldots, s_n$ and is bound to maintain consistency and accept their (accepted) deductive consequences. In fact, she would be in dereliction of her duty as theorist if she accepted the preface sentence in the body of the book. However, the context of the preface is different: here it is customary to drop the professional exigencies of the theorist and acknowledge broader features of the author's cognitive life. She has fulfilled her duty as theorist and done the utmost to accept only those claims that are justified by her evidence and arguments. However, some of these conclusions may not yet be attended with the inner glow of belief. Perhaps, if the work meets with no devastating objections, she may eventually cease to believe the humble claim in the preface. Thus the distinction between belief and acceptance explains why we are not alarmed by the sentence in the context of the preface, but we would be shocked if we saw it used as a premise in the body of the text. For a similar resolution of the paradox, see Chapter 5 of Stalnaker (1984).

It is easy to underestimate the consequences of accepting Cohen's arguments. For one, we would have to reinterpret all of the theories of rational belief that we have discussed as theories of rational acceptance. In fact, there may be no theory of rational belief, but only psychological tricks and heuristics for coming to believe, similar to those Pascal recommends for arriving at faith in Christ. Longstanding dogmas about the relation between belief and knowledge would have to be revisited. Moreover, excessive appeal to the distinction threatens the unity and cohesiveness of our cognitive lives. For a discussion of these kinds of objections see Kvanvig (2016). For an overview of the distinction between belief and acceptance, see Weirich (2004).

3 STRUCTURES FOR PARTIAL BELIEF

3.1 *Bayesianism*

Bayesianism, or subjective probability theory, is by far the dominant paradigm for modeling partial belief. The literature on the subject is by now very large and includes many approachable introductions. The summary provided here will, of necessity, be rather brief. For an article-length introduction see Huber (2016), Easwaran (2011a, 2011b), or Weisberg (2011). For a book-length introduction see Earman (1992) or Howson and Urbach (2006). For an article-length introduction to Bayesian models of rational action, see Briggs (2017) or Thoma (this volume). For an approachable book-length introduction to the theory of rational choice see Resnik (1987).

The heart of the Bayesian theory is roughly the following:

1. There is a fundamental psychological attitude called *degree of belief* (sometimes called *confidence* or *credence*) that can be represented by numbers in the $[0, 1]$ interval.

2. The degrees of belief of rational agents satisfy the axioms of probability theory.

3. The degrees of belief of rational agents are *updated* by some flavor of probabilistic conditioning.

The first two principles are the synchronic requirements of Bayesian theory; the third principle concerns diachronic updating behavior. Most Bayesians would also agree to some version of the following principles, which link subjective probabilities with deliberation and action:

4. Possible states of the world (sometimes *outcomes*) are assigned a *utility*: a positive or negative real number that reflects the desirability or undesirability of that outcome.

5. Rational agents perform only those actions that maximize *expected* utility, which is calculated by weighing the utility of outcomes by their subjective probability.

What makes Bayesianism so formidable is that, in addition to providing an account of rational belief and its updating, it also provides an account of rational action and deliberation. No other theory can claim a developed, fine-grained account of all three of these aspects of belief. In the following we will briefly spell out some of the technical details of the Bayesian picture.

3.1.1 *Probabilism*

In this section we flesh out the details of the synchronic component of the Bayesian theory. For the purposes of this section we will take propositions to be the objects of (partial) belief. It is also possible to take a syntactic approach and assign degrees of belief to sentences in a formal language. For the most part, nothing hinges on which approach we choose. For arguments in favor of the syntactic approach, see Weisberg (2011).

As usual, let W be a set of possible worlds. Let $\mathcal{F}$ be a σ-field over W.[3] A *credence* function p assigns a degree of belief to every proposition in $\mathcal{F}$. *Probabilism* requires that the credence function satisfies the axioms of probability. For every $E, F \in \mathcal{F}$:

$$p(E) \text{ is a positive, real number;} \tag{Positivity}$$

$$p(W) = 1; \tag{Unitarity}$$

$$\text{if } E \cap F = \varnothing, \text{ then } p(E \cup F) = p(E) + p(F). \tag{Additivity}$$

From these principles it is possible to derive several illuminating theorems. For example, the degree of belief assigned to the contradictory proposition is equal to zero. Furthermore, if E entails F, then $p(E) \leq p(F)$. Finally, for any proposition $E \in \mathcal{F}$ we have that $0 \leq p(E) \leq 1$.

In the standard axiomatization of probability theory due to Kolmogorov (1950), additivity is strengthened to Countable Additivity.

$$\begin{aligned}&\text{If } E_1, E_2, \ldots \text{ are mutually exclusive,}\\ &\qquad \text{then } p(\textstyle\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i).\end{aligned} \tag{Countable Additivity}$$

This requirement is not as innocent as it looks: it rules out the possibility that any agent is indifferent over a countably infinite set of mutually exclusive possibilities. De Finetti (1970, 1972) famously argued that we ought to reject countable additivity since it is conceivable that God could pick out a natural number "at random" and with equal (zero) probability. For another example, suppose you assign 50% credence to the proposition $\neg B$ that not all ravens that will ever be observed are black. Let $\neg B_i$ be the proposition that the i^{th} observed raven is the first non-black raven to appear. Then $\neg B = \bigcup_{i=1}^{\infty} \neg B_i$. Countable additivity entails that for all $\epsilon > 0$ there is a finite n such that $p(\bigcup_{i=1}^{n} \neg B_i) = 1/2 - \epsilon$. So you must be nearly certain that if all ravens are not black, the first non-black raven will appear among the first n ravens. The only way to assign equal probability to all $\neg B_i$ is to violate countable additivity by setting $p(\neg B_i) = 0$ for all i. This solution has its own drawbacks. On all standard models of Bayesian update it will be impossible to become convinced that the i^{th} raven is

3 Refer to Section 1 if you need to refresh yourself on the definition of a σ-field.

indeed non-black, even if you are looking at a white one. For more on countable additivity, see Chapter 13 in Kelly (1996).

Now that we have defined probabilism, it is natural to ask how to justify it: why *should* a rational agent's degrees of belief obey the probability axioms? There are roughly two kinds of answers to this question current in the Bayesian canon.

The traditional answer is that an agent that violates the axioms of probability opens herself up to systems of bets that, although fair from the agent's perspective, guarantee a sure loss. Answers of this flavor are called *Dutch book* arguments and they require positing some connection between degrees of belief and fair betting quotients. Some epistemologists find Dutch book arguments to be unconvincing either because they disavow any tight connection between degrees of belief and betting quotients, or they deny that any facts about something so pragmatic as betting could have normative epistemic force. These epistemologists tend to prefer *accuracy* arguments, which purport to show that any agent violating the probability axioms will have beliefs which are less accurate, or "further from the truth," than agents that satisfy the axioms. We will briefly review the traditional Dutch book-style arguments. For the original articulation of the accuracy perspective see Joyce (1998). For an article-length overview of accuracy-style arguments see Pettigrew (2016b). For a book-length treatment see Pettigrew (2016a).

Dutch book arguments require specifying some connection between degrees of belief and fair betting quotients. For de Finetti (1937) the connection was definitional: an agent's degree of belief in a proposition A simply is her fair odds ratio for a bet that pays \$1 if A is true and nothing otherwise. If you are willing to pay at most \$.50 for a bet that pays \$1 if A is true and nothing otherwise, then it must be that your degree of confidence in A is 50%. It is easy to see what is wrong with this kind of definition: there may be factors other than the subject's degree of belief which affect her fair betting quotient. She may be risk averse, or risk-loving; she may abhor gambling, or love showing off. Ramsey (1931) avoids some of these problems by pricing bets in utility, rather than money, and appealing to the existence of an "ethically neutral" proposition that is considered equally likely to be true and false. For more on the connection between degrees of belief and betting ratios see Eriksson and Hájek (2007).

Supposing that a suitable connection between degrees of belief and fair betting quotients exists, it is possible to construct a "Dutch book" against an agent violating the axioms of probability. To get such an argument going we suppose that if the agent's degree of belief in A is $p(A)$, then she considers fair a bet that costs $\$p(A) \cdot Y$ and pays $\$Y$ if A is true and \$0 otherwise. Note that we allow Y to take positive and negative values. This means that the agent is willing to assume the role of the bookie and sell a bet that "costs" $-\$p(A) \cdot Y$ and "pays" $-\$Y$ if A is true and \$0 otherwise. Now suppose that such an agent violates finite additivity. One way this may happen is if for A, B such that $A \cap B = \varnothing$, we have that $p(A \cup B) > p(A) + p(B)$. Then, the agent considers fair

1. a bet that costs $-\$p(A)$ and pays $-\$1$ if A is true and \$0 otherwise;
2. a bet that costs $-\$p(B)$ and pays $-\$1$ if B is true and \$0 otherwise;
3. a bet that costs $\$p(A \cup B)$ and pays \$1 if $A \cup B$ is true and \$0 otherwise.

There are three possible scenarios: either A and B are both false or exactly one of them is true. The reader should confirm that in any of these scenarios the agent is left with exactly $\$p(A) + \$p(B) - \$p(A \cup B) < 0$. By reversing which bets the agent buys and sells, we can construct a Dutch book against an agent that violates additivity by having $p(A \cup B) < p(A) + p(B)$. Similar strategies work to construct Dutch books against agents that violate Positivity, Unitarity, and Countable Additivity. Furthermore, it is possible to show that if your degrees of belief validate the probability axioms, then no Dutch book can be made against you (Kemeny, 1955). For more on Dutch book arguments see Section 3.3 in Hájek (2012).

3.1.2 *Updating by Conditioning*

We have discussed the synchronic content of the Bayesian theory, but we still need to talk about how degrees of belief are updated upon receiving new information. There are two standard models of partial belief update: strict conditionalization and Jeffrey conditionalization. Strict conditionalization assumes that the information received acquires the maximal degree of belief. Jeffrey conditionalization allows for the situation in which no proposition is upgraded to full certainty when new information is acquired.

For all propositions $A, B \in \mathcal{F}$ such that $p(A) > 0$, the conditional probability of B given A is defined as:

$$p(B \mid A) := \frac{p(A \cap B)}{p(A)}.$$

Conditionalization by A restricts all possibilities to those compatible with A and then renormalizes by the probability of A to ensure that unitarity holds. By far the most standard modeling of partial belief update holds that degrees of belief ought to be updated by conditionalization. In other words, if p_t is your credence function at time t and A is a proposition expressing the total new information acquired by $t' > t$, then $p_{t'}$ ought to equal $p_t(\cdot \mid A)$, whenever $p_t(A) > 0$.

What if new information does not render any proposition certain, but merely changes the subjective probability of some propositions? Jeffrey (1983) proposes the following update rule. Suppose that p_t is your credence function at time t. Suppose that the total evidence that comes in by time t' updates your degrees of belief in partition $\{A_i\}_{1\leq i\leq n}$ (and in no finer partition) setting each respectively to a_i with $\sum_i a_i = 1$. Then your new credence function $p_{t'}$ ought to be $\sum_i p(\cdot \mid A_i)a_i$.

Why should a rational agent update by strict or Jeffrey conditionalization? Dutch-book style arguments for strict conditionalization are given in Teller (1973) and Lewis (1999) and extended to Jeffrey conditionalization in Armendt (1980). For more see Skyrms (2009). For an accuracy-style argument in favor of strict conditionalization and against Jeffrey conditionalization, see Leitgeb and Pettigrew (2010).

For our purposes it is important to point out that conditional probability is always a lower bound for the probability of the material conditional. In other words, for all $E, H \in \mathcal{F}$,

$$p(H \mid E) \leq p(E \rightarrow H),$$

whenever $p(E) > 0$. We can see this as a quantitative version of the qualitative principle of Conditionalization we discussed in Section 2.1.1: however confident a Bayesian agent becomes in H after updating on E, she must have been at least as confident that H is a material consequence of E. Popper and Miller (1983) took this observation to be "completely devastating to the inductive interpretation of the calculus of probability." For the history of the Popper-Miller debate see Chapter 4 in Earman (1992). A similar property can be demonstrated for Jeffrey conditioning (Genin, 2017).

Both strict and Jeffrey conditionalization are defined in terms of conditional probability. The probability of B conditional on A is standardly defined as the ratio of the unconditional probabilities $p(A \cap B)$ and $p(A)$. Clearly, this ratio is undefined when $p(A) = 0$. Some theorists would like conditional probability to be defined even when conditioning on propositions of probability zero. The standard approach in mathematical statistics, due to Kolmogorov (1950), is via the *conditional expectation*. On that approach, conditional probability remains dependent on unconditional probability. An alternative approach, adopted by Popper (1955) and Renyi (1955), takes conditional probability as a primitive, rather than a derivative, notion. For a defense of the conditional expectation, see Gyenis, Hofer-Szabó, and Rédei (2017). For an introduction to primitive conditional probabilities, see Easwaran (this volume). For a critique of the standard notion of conditional probability, see Hájek (2003).

	States	
	Good	Rotten
Acts	**Outcomes**	
Break into Bowl	6-Egg Omelet	No Omelet
	+10	0
Break into Saucer	6-Egg Omelet, extra washing	5-Egg Omelet, extra washing
	+8	+4

Table 1: A payoff table for the morning chef

3.1.3 *Deliberation and Action*

One of the signal advantages of the Bayesian model of partial belief is that it is ready-made to plug into a prominent model of practical deliberation. Decision theory, or rational choice theory, is too large and sprawling a subject to be effectively covered here, although it will be presented in cursory outline. For an excellent introduction, see Briggs (2017) or Thoma (this volume). For our purposes, it is enough to note that a well-developed theory exists and that no comparable theory exists for alternative models of belief.[4]

Suppose you would like to make a six egg omelet. You've broken 5 fresh eggs into a mixing bowl. Rooting around your fridge, you find a loose egg of uncertain provenance. If you are feeling lucky you can break the suspect egg directly into the mixing bowl; if you are wary of the egg, you might break it into a saucer first and incur more dishwashing.

There are four essential ingredients to this sort of decision-theoretic situation. There are *outcomes*, over which we have defined *utilities* measuring the desirability of the outcome. In the case of the omelet the outcomes are a ruined omelet or a 5–6 egg omelet, with or without extra washing. There are *states*—usually unknown to and out of the control of the actor—which influence the outcome of the decision. In our case the states are exhausted by the possible states of the suspect egg: either good or rotten. Finally, there are *acts* which are under the control of the decision maker. In our case the acts include breaking the egg into the bowl or the saucer. Of course there are other conceivable acts: you might throw the suspect egg away and make do with a 5-egg omelet; you might even flip a coin to decide what to do. We omit these for the sake of simplicity. These four elements are usually summarized in a payoff table (see Table 1). To fit this into the

4 However, recent work such as Lin (2013) and Spohn (2017, 2019) may remedy that inadequacy in the case of qualitative belief.

framework of partial belief we assume that the set of acts $A_1, A_2, \ldots, A_n$ partition W. We also assume the set of states $S_1, S_2, \ldots, S_m$ partition W. We assume that the credence function assigns a probability to every outcome. We assume that acts and states are logically independent, so that no state rules out the performance of any act. Finally, we assume that given a state of the world S_j and an act A_i there is exactly one outcome O_{ij}, which is assigned a utility $U(O_{ij})$. The ultimate counsel of rational choice theory is that agents ought to perform only those acts that maximize *expected* utility. The expected utility of an act is defined as:

$$EU(A_i) = \sum_{j=1}^{m} p_{A_i}(S_j)U(O_{ij}),$$

where $p_{A_i}(S_j)$ is roughly how likely the agent considers S_j given that she has performed act A_i. Difficulties about how this quantity should be defined give rise to the schism between evidential and causal decision theory (see Section 3.3 in Thoma, this volume). However, in many situations, including the dilemma of the omelet, the act chosen does not affect the probabilities with which states obtain. This is called "act-state independence" in the jargon of rational choice theory. In cases of act-state independence there is broad consensus that $p_{A_i}(S_j)$ should be equal to the unconditional degree of belief $p(S_j)$.

Central to the literature on decision theory are a number of *representation theorems* showing that every agent with qualitative preferences satisfying a set of rationality postulates can be represented as an expected utility maximizer (von Neumann & Morgenstern, 1944; Savage, 1954). These axioms are controversial, and are subject to intuitive counterexamples. Allais (1953) and Ellsberg (1961) give examples in which seemingly rational agents violate the rationality postulates and therefore cannot, even in principle, be represented as expected utility maximizers. For more on this subject, see Sections 2 and 3 in Briggs (2017).

3.1.4 *Modifications and Alternatives*

Dissatisfaction with various aspects of the Bayesian theory has spawned a number of formal projects. Many epistemologists reject the notion that rational agents must have *precise* credences in every proposition that they can entertain; instead they claim that rational agents may have *imprecise* credences representable by *intervals* of real numbers. For an introduction to imprecise probability, see Mahtani (this volume). The theory of Dempster-Shafer belief functions (Dempster, 1968; Shafer, 1976) rejects the tight connection between degrees of belief and fair betting ratios. Fair betting ratios ought indeed satisfy the axioms of probability, but degrees of belief need not. Nevertheless, it should be possible to calculate fair betting ratios from degrees of belief when these are necessary. For this purpose, degrees of belief may satisfy a weaker set of axioms than those of the probability calculus. For an introduction to Dempster-Shafer belief functions see Section 3.1 in Huber (2016).

Many epistemologists have held that degrees of belief are not so definitely comparable as suggested by the probabilistic analysis. Keynes (1921) famously proposes that degrees of belief may enjoy only an *ordinal* structure, which admits of qualitative, but not quantitative, comparison. Keynes even suggests that the strength of some pairs of partial beliefs cannot be compared at all. Koopman (1940) and Fine (1973) pursue Keynes' suggestions, developing axiomatic theories of qualitative probability. See Konek (this volume) for an introduction to qualitative probability comparisons.

3.2 *Ranking Theory*

Cohen (1977, 1980) distinguishes between two rival probabilistic traditions. Pascalian probability finds its latest expression in contemporary Bayesianism. But Cohen traces a rival tradition back to Francis Bacon. Roughly, these two can be distinguished by the *scale* they select for the strength of belief. On the Pascalian scale a degree of belief of zero in some proposition implies maximal conviction in its negation. On the Baconian scale, a degree of belief of zero implies no conviction in either the proposition or its negation. Thus, the Pascalian scale runs from "disproof to proof" whereas the Baconian runs from "no evidence, or non-proof to proof" (Cohen, 1980, p. 224). Cohen (1977) argues that despite the conspicuous successes of Pascalian probability, the Baconian scale is more appropriate in other settings, including legal proceedings.

Ranking theory, first developed in Spohn (1988), is a sophisticated contemporary theory of Baconian probability. For an article-length introduction to ranking theory see Huber (2013; this volume). For an extensive book-length treatment, with applications to many subjects in epistemology and philosophy of science, see Spohn (2012). We mention some of its basic features, as it provides a useful counterpoint to the models of belief we have already discussed.

As before, let W be a set of possible worlds. Let $\mathcal{F}$ be an algebra over W.[5] A function $\beta : \mathcal{F} \to \mathbb{N} \cup \{\infty\}$ from $\mathcal{F}$ into the set of natural numbers $\mathbb{N}$ extended by ∞, is a *positive ranking function* on $\mathcal{F}$ just in case for any $A, B \in \mathcal{F}$:

$$\beta(\varnothing) = 0; \qquad \text{(Consistency)}$$

$$\beta(W) = \infty; \qquad \text{(Infinitivity)}$$

5 Refer to Section 1 if you need to refresh yourself on the definition of an algebra

$$\beta(A \cap B) = \min\{\beta(A), \beta(B)\}. \qquad \text{(Minimitivity)}$$

A positive ranking function expresses degrees of belief. If $\beta(A) > 0$, then we may say that A is (fully) believed and $\neg A$ is disbelieved. If $\beta(A) = 0$ then A is not believed and $\neg A$ may not be believed either. Thus, ranking theory can be seen as satisfying the "Lockean thesis," the intuitive proposal that a degree of belief above some threshold is necessary and sufficient for full belief (see Section 5.2). Note however that nothing in ranking theory requires us to say that the threshold is exactly zero: we could have chosen any positive number n.

Let β be a positive ranking function and $A \in \mathcal{F}$ with $\beta(\neg A) < \infty$. Then for any $B \in \mathcal{F}$ the *conditional positive rank of B given A* is defined as

$$\beta(B \mid A) = \beta(\neg A \cup B) - \beta(\neg A).$$

The function $\beta_A : B \mapsto \beta(B \mid A)$ is called the *conditionalization of β by A* and is itself a positive ranking function. This definition is used to articulate an update rule for ranking theory: if β is your positive ranking function at time t and between t and t' you become certain of $E \in \mathcal{F}$ and no logically stronger proposition, then β_E should be your new ranking function at time t'. Spohn (1988) also defines ranking-theoretic analogues of Jeffrey conditioning.

It is clear from the definition of conditioning that, as in the Bayesian case, the rank of the material conditional is a lower bound for the conditional rank: $\beta(A \to B) \leq \beta(B \mid A)$. It also satisfies a version of Rational Monotony: if $\beta(\neg A) = 0$ and $\beta(B) > 0$, then $\beta(B \mid A) > 0$.[6] Therefore, ranking-theoretic update satisfies the "spirit" of AGM update. Note however, that ranking theory has no trouble with iterated belief revision: a revision takes as input a ranking function and an evidential proposition and outputs a new ranking function.

Ranking theory lies somewhat awkwardly between a theory of full and partial belief. On the one hand, all propositions of positive rank are fully believed. On the other hand, the rank of a proposition measures something about the strength of that belief. But how should we interpret these ranks? Huber (this volume) investigates the relation between ranking-theoretic degrees of belief, and AGM-style degrees of entrenchment. The *degree of entrenchment* for a proposition A is defined as the number of independent and reliable information sources testifying *against* A that it requires for the agent to give up full belief in A. Degrees of entrenchment may be used to *measure* ranking-theoretic degrees of belief; alternatively, it is possible to *identify* ranking-theoretic degrees of belief with degrees of entrenchment. Huber (manuscript) proves that if an agent defines her full beliefs from

6 Rational Monotony is not satisfied if we set the the threshold for full belief at some number greater than zero (Raidl, forthcoming, footnote 26).

an entrenchment function, her beliefs will be consistent and deductively closed iff the entrenchment function is a ranking function.

One of the advantages of ranking theory over AGM is that it allows *reasons* to be defined (Spohn, 2012). Say that *A is a reason for B* with respect to the positive ranking function β iff $\beta(B \mid A) > \beta(B \mid \neg A)$. Say that an agent *has A as a reason for B* iff A is a reason for B according to her positive ranking function β and $\beta(A) > 0$. Note that it is not possible to make such a definition in the AGM theory since the conditional degree of entrenchment is not defined. Thus ranking theory may provide an answer to Pollock's criticism of belief revision by allowing various kinds of defeat of reasons to be represented (Spohn, 2012, Section 11.5).

4 ELIMINATIONISMS

There are those who deny that there are any interesting principles bridging full and partial belief. Theorists of this persuasion often want either to eliminate one of these attitudes or reduce it to a special case of the other. Jeffrey (1970) suggests that talk of full belief is vestigial and will be entirely superseded by talk of partial belief and utility:

> ...nor am I disturbed by the fact that our ordinary notion of *belief* is only vestigially present in the notion of degree of belief. I am inclined to think Ramsey sucked the marrow out of the ordinary notion, and used it to nourish a more adequate view. But maybe there is more there, of value. I hope so. Show me; I have not seen it at all clearly, but it may be there for all that (p. 172).

Theorists such as Kaplan (1996) also suggests that talk of full belief is superfluous once the mechanisms of Bayesian decision theory are in place. After all, only partial beliefs (or *confidence* in Kaplan's terminology) and utilities play any role in the Bayesian framework of rational deliberation, whereas full belief need not be mentioned at all. Those committed to full beliefs have the burden of showing what difference they make to our lives:

> Making the case that talk of investing confidence leaves out something important—something we have in mind when we talk of belief—is going to require honest toil. One has to say ...exactly how an account of rational human activity will be the poorer if it has no recourse to talk of belief. In short, one has to meet *the Bayesian Challenge*. (p. 100)

Stalnaker (1984) is much more sympathetic to a qualitative notion of belief (or acceptance) but acknowledges the force of the Bayesian Challenge.

> Bayesian decision theory gives a complete account of how probability values . . . ought to guide behavior . . . So what could be the point of selecting an interval near the top of the probability scale and conferring on the propositions whose probability falls in that interval the honorific title "accepted"? Unless acceptance . . . makes a difference to how the agent behaves, or ought to behave it is difficult to see how the concept of acceptance can have the interest and importance for inquiry that it seems to have. (p. 91)

It is true that there is no canonical qualitative analogue to the Bayesian theory of practical deliberation. However, the fact that it is the theorist of full belief that feels the challenge, and not *vice versa*, may be an accident of history: if a qualitative theory of practical deliberation had been developed first, the shoe would now be on the other foot. The situation would be even more severe if qualitative decision making, which we seem to implement as a matter of course, were less cognitively demanding than its Bayesian counterpart. Of course, this anticipates a robust theory of rational qualitative deliberation that is not immediately forthcoming. However, recent work such as Lin (2013) and Spohn (2017, 2019) may remedy that inadequacy. For example, Lin (2013) proves a Savage-style representation theorem characterizing the relationship between full beliefs, desires over possible outcomes, and preferences over acts. By developing a theory of rational action in terms of qualitative belief, Lin demonstrates how one might answer the Bayesian challenge.

On the other hand there are partisans of full belief that are deeply skeptical about partial beliefs.[7] Many of these object that partial beliefs have no psychological reality and would be too difficult to reason with if they did. Horgan (2017) goes so far as to say that typically "there is no such psychological state as the agent's credence in p" and that Bayesian epistemology is "like alchemy and phlogiston theory: it is not about any real phenomena, and thus it also is not about any genuine norms that govern real phenomena" (p. 7). Harman (1986) argues that we have very few explicit partial beliefs. A theory of reasoning, according to Harman, can concern only explicit attitudes, since these are the only ones that can figure in a reasoning process. Therefore, Bayesian epistemology, while perhaps an account of dispositions to act, is not a guide to reasoning. Nevertheless, partial beliefs may be implicit in our system of full beliefs in that they can be reconstructed from our dispositions to revise them.

> How should we account for the varying strengths of explicit beliefs? I am inclined to suppose that these varying strengths are

7 See Harman (1986), Pollock (2006), Moon (2017), and Horgan (2017). See also the "bad cop" in Hájek and Lin (2017).

> implicit in a system of beliefs one accepts in a yes/no fashion. My guess is that they are to be explained as a kind of epiphenomenon resulting from the operation of rules of revision. For example, it may be that P is believed more strongly than Q if it would be harder to stop believing P than to stop believing Q, perhaps because it would require more of a revision of one's view... (Harman, 1986, p. 22)

On this picture, almost all of our explicit beliefs are qualitative. Partial beliefs are not graded *belief attitudes* toward propositions, but rather dispositions to revise our *full* beliefs. The correct theory of partial belief, according to Harman, has more to do with entrenchment orders (see Section 2.2.2) or ranking-theoretic degrees of belief (see Section 3.2) than with probabilities. Other apparently partial belief attitudes are explained as full beliefs *about* objective probabilities. So, in the case of a fair lottery with ten thousand tickets, the agent does not believe to a *high degree* that the n^{th} ticket will not win, but rather fully believes that it is objectively improbable that it will win.

Frankish (2009) objects that Harman's view requires that an agent have a full belief in any proposition that we have a degree of belief in: "And this is surely wrong. I have some degree of confidence (less than 50%) in the proposition that it will rain tomorrow, but I do not believe flat-out that it will rain—not, at least, by the everyday standards for flat-out belief" (p. 4). Harman might reply that Frankish merely has a full belief in the objective probability of rain tomorrow. Frankish claims that this escape route is closed to Harman because single events "do not have objective probabilities," but this matter is hardly settled.

Staffel (2013) gives an example in which a proposition with a higher degree of belief is apparently less entrenched than one with a lower degree of belief. Suppose that you will draw a sequence of two million marbles from a big jar full of red and black marbles. You do not know what proportion of the marbles are red. Consider the following cases.

SCENARIO 1. You have drawn twenty marbles, 19 black and one red. Your degree of belief that the last marble you will draw is black is .95.

SCENARIO 2. You have drawn a million marbles, 900,000 of which have been black. Your degree of belief that the last marble you will draw is black is $19/20 = .90$.

Staffel argues that your degree of belief in the first case is higher than in the second, but much more entrenched in the second than in the first. Therefore, degree of belief cannot be reduced to degree of entrenchment. Nevertheless, the same gambit is open to Harman in the case of the marbles—he can claim that in both scenarios you merely have a full belief in a proposition about objective chance. See Staffel (2013) for a much more extensive engagement with Harman (1986).

5 BRIDGE PRINCIPLES FOR FULL AND PARTIAL BELIEF

Anyone who allows for the existence of both full and partial belief inherits a thorny problem: how are full beliefs related to partial beliefs? That seemingly innocent question leads to a treacherous search for *bridge principles* connecting a rational agent's partial beliefs with her full beliefs. Theorists engaged in the search for bridge principles usually take for granted some set of rationality principles governing full belief and its revision e.g. AGM theory, or a rival system of non-monotonic reasoning. Theorists usually also take for granted that partial belief ought to be representable by probability functions obeying some flavor of Bayesian rationality. The challenge is to propose additional rationality postulates governing how a rational agent's partial beliefs cohere with her full beliefs. In this section, we will for the most part accept received wisdom and assume that orthodox Bayesianism is the correct model of partial belief and its updating. We will be more open-minded about the modeling of full belief and its rational revision.

In this section, we will once again take propositions to be the objects of belief. In the background, there will be a (usually finite) set W of possible worlds. As before, the reader is invited to think of W as a set of coarse-grained, mutually exclusive, possible ways the actual world might be. The actual world is assumed to instantiate one of these coarse-grained possibilities. We write $\mathcal{B}$ to denote the set of *propositions* that the agent believes and use $\mathcal{B}(A)$ as shorthand for $A \in \mathcal{B}$. We will also require some notation for qualitative propositional belief change. For all $E \subseteq W$, write $\mathcal{B}_E$ for the set of propositions the agent would believe upon learning E and no stronger proposition. We will also write $\mathcal{B}(A \mid E)$ as shorthand for $A \in \mathcal{B}_E$. By convention, $\mathcal{B} = \mathcal{B}_W$. If $\mathcal{F}$ is a set of propositions, we let $\mathcal{B}_{\mathcal{F}}$ be the set $\{\mathcal{B}_E : E \in \mathcal{F}\}$. The set $\mathcal{B}_{\mathcal{F}}$ represents an agent's *dispositions* to update her qualitative beliefs given information from $\mathcal{F}$.

The following normative constraint on the set of full beliefs $\mathcal{B}$ plays a large role in what follows.

For all propositions $A, B \subseteq W$: (Deductive Cogency)

1. $\mathcal{B}(W)$;
2. not $\mathcal{B}(\varnothing)$;
3. if $\mathcal{B}(A)$ and $A \subseteq B$, then $\mathcal{B}(B)$;
4. if $\mathcal{B}(A)$ and $\mathcal{B}(B)$ then $\mathcal{B}(A \cap B)$.

The first two clauses say that the agent believes the true world to be among the worlds in W and that she does not believe the empty set to contain the true world. The third clause says that belief is closed under single-premise entailment, i.e. if the agent believes A and A logically entails B, then she believes B. The final clause says that the agent's beliefs are closed under conjunction, i.e. if she believes A and she believes B, then she believes $A \cap B$. Together, clauses 3 and 4 say that the agent's beliefs are closed under entailment by finitely many premises. When W is finite, the set $\mathcal{B}$ must be finite as well, implying that Deductive Cogency is equivalent to the following formulation:

$$\mathcal{B} \text{ is consistent and } \mathcal{B}(B) \text{ iff } \cap\mathcal{B} \subseteq B. \qquad \text{(Deductive Cogency)}$$

In other words, Deductive Cogency means that there is a single, non-empty proposition, which is the logically strongest proposition that the agent believes, entailing all her other beliefs. When the two formulations of Deductive Cogency come apart, we will always mean the latter one. Deductive Cogency only mentions the set of full beliefs $\mathcal{B}$, and is therefore not a bridge principle at all. Bridge principles are expressed as constraints holding for pairs $\langle \mathcal{B}, p \rangle$.

All of the rationality norms that we have seen for updating qualitative beliefs have propositional analogues. The following are propositional analogues for the six basic AGM principles. Here E, F are arbitrary subsets of W.

$$\mathcal{B}_E = Cn(\mathcal{B}_E). \qquad \text{(Closure)}$$

$$E \in \mathcal{B}_E. \qquad \text{(Success)}$$

$$\mathcal{B}_E \subseteq Cn(\mathcal{B} \cup \{E\}). \qquad \text{(Inclusion)}$$

$$\text{If } \neg E \notin Cn(\mathcal{B}) \text{ then } \mathcal{B} \subseteq \mathcal{B}_E. \qquad \text{(Preservation)}$$

$$\mathcal{B}_E \text{ is consistent if } E \neq \varnothing. \qquad \text{(Consistency)}$$

$$\text{If } E \equiv F\text{, then } \mathcal{B}_E = \mathcal{B}_F. \qquad \text{(Extensionality)}$$

$$\mathcal{B}_{E \cap F} \subseteq Cn(\mathcal{B}_E \cup \{F\}). \qquad \text{(Conjunctive inclusion)}$$

$$\text{If } \neg F \notin Cn(\mathcal{B}_E)\text{, then } Cn(B_E \cup \{F\}) \subseteq B_{E \cap F}. \qquad \text{(Conjunctive preservation)}$$

Supposing that for all $E \subseteq W$, $\mathcal{B}_E$ satisfies Deductive Cogency, the first six postulates reduce to the following three, for arbitrary $E \subseteq W$.

$$\cap\mathcal{B}_E \subseteq E. \qquad \text{(Success)}$$

$$\cap\mathcal{B} \cap E \subseteq \cap\mathcal{B}_E. \qquad \text{(Inclusion)}$$

$$\text{If } \cap\mathcal{B} \not\subseteq \neg E, \text{ then } \cap\mathcal{B}_E \subseteq \cap\mathcal{B} \cap E. \quad \text{(Preservation)}$$

Together, Inclusion and Preservation say that whenever information E is consistent with current belief $\cap\mathcal{B}$,

$$\cap\mathcal{B}_E = \cap\mathcal{B} \cap E.$$

If $\mathcal{F}$ is a collection of propositions and for all $E \in \mathcal{F}$, the belief sets $\mathcal{B}, \mathcal{B}_E$ satisfy the AGM principles, we say that $\mathcal{B}_{\mathcal{F}}$, the agent's disposition to update her qualitative beliefs given information from $\mathcal{F}$, satisfies the basic AGM principles.

We will use $p(\cdot)$ to denote the probability function representing the agent's partial beliefs. Of course, $p(\cdot)$ is defined on a σ-algebra of subsets of W. In the usual case, when W is finite, we can take the $\wp(W)$ to be the relevant σ-algebra. To update partial belief, we adopt the standard probabilistic modeling. For $E \subseteq W$ such that $p(E) > 0$, $p(\cdot \mid E)$ is the partial belief function resulting from learning E. We will sometimes use p_E as a shorthand for $p(\cdot \mid E)$. Almost always, partial belief is updated via conditioning:

$$p(A \mid E) = \frac{p(A \cap E)}{p(E)}, \text{ whenever } p(E) > 0.$$

Let $\mathcal{F}_p^+$ be the set of propositions with positive probability according to p, i.e $\{A \subseteq W : p(A) > 0\}$.

5.1 *Belief as Extremal Probability*

The first bridge principle that suggests itself is that full belief is just the maximum degree of partial belief. Expressed probabilistically, it says that at all times a rational agent's beliefs and partial beliefs can be represented by a pair $\langle \mathcal{B}, p \rangle$ satisfying:

$$\mathcal{B}(A) \text{ iff } p(A) = 1. \quad \text{(Extremal Probability)}$$

Roorda (1995) calls this the *received view* of how full and partial belief ought to interact. Gärdenfors (1986) is a representative of this view, as are van Fraassen (1995) and Arló-Costa (1999), although the latter two accept a slightly non-standard probabilistic modeling for partial belief. For fans of Deductive Cogency, the following observations ought to count in favor of the received view.

Theorem 11 *If $\langle \mathcal{B}, p \rangle$ satisfy extremal probability, then $\mathcal{B}$ is deductively cogent.*

Gärdenfors (1986) proves the following.

Theorem 12 *Suppose that $\langle \mathcal{B}_E, p_E \rangle$ satisfy extremal probability for all $E \in \mathcal{F}_p^+$. Then $\mathcal{B}_{\mathcal{F}_p^+}$ satisfies the AGM postulates.*

In other words: if an agent's partial beliefs validate the probability axioms, she updates by Bayesian conditioning and fully believes all and only those propositions with extremal probability, her qualitative update behavior will satisfy all the AGM postulates (at least whenever Bayesian conditioning is defined). Readers who take the AGM revision postulates to be a *sine qua non* of rational belief update will take this to be good news for the received view.

Roorda (1995) makes three criticisms of the received view. Consider the following three propositions.

1. Millard Fillmore was the 13th President of the United States;

2. Millard Fillmore was a U.S. President;

3. Millard Fillmore either was or was not a U.S. President.

Of course, I am not as confident that Fillmore was the 13^{th} president as I am in the truth of the tautology expressed in (3). Yet there does not seem to be anything wrong with saying that I fully believe each of (1), (2), and (3). However, if extremal probability is right, it is irrational to fully believe each of (1), (2), and (3) and not assign them all the same degree of belief.

Roorda's second objection appeals to the standard connection between degrees of belief and practical decision making. Suppose I fully believe (1). According to the standard interpretation of degrees of belief in terms of betting quotients, I ought to be accept a bet that pays out a dollar if (1) is true, and costs me a million dollars if (1) is false. In fact, if I truly assign unit probability to (1), I ought to accept nearly any stakes whatsoever that guarantee some positive payout if (1) is true. Yet it seems perfectly rational to fully believe (1) and refrain from accepting such a bet. If we accept Bayesian decision theory, extremal probability seems to commit me to all sorts of weird and seemingly irrational betting behavior.

Roorda's final challenge to extremal probability appeals to *corrigibility*, according to which it is reasonable to believe that at least some of my beliefs may need to be abandoned in light of new information. However, if partial beliefs are updated via Bayesian conditioning, I can never cease to believe any of my full beliefs since if $p(A) = 1$ it follows that $p(A \mid E) = 1$ for all E such that $p(E) > 0$. If we believe in Bayesian conditioning, extremal probability seems to entail that I cannot revise any of my full beliefs in light of new information.

5.2 *The Lockean Threshold*

The natural response to the difficulties with the received view is to retreat from full certainty. Perhaps full belief corresponds to partial belief above some *threshold* falling short of certainty. Foley (1993) dubbed this view the *Lockean thesis*, after some apparently similar remarks in Book IV of Locke's *Essay Concerning Human Understanding*. So far, the Lockean thesis is actually ambiguous. There may be a single threshold that is rationally mandated for all agents and in all circumstances. Alternatively, each agent may have her own threshold that she applies in all circumstances—that threshold may characterize how "bold" or "risk-seeking" the agent is in forming qualitative beliefs. A yet weaker thesis holds that the threshold may be contextually determined. We distinguish the strong, context-independent Lockean thesis (SLT) from the weaker, context-dependent thesis (WLT). The domain of the quantifier may be taken as the set of all belief states $\langle \mathcal{B}, p \rangle$ a *particular* agent may find herself in, or as the set of all belief states whatsoever.

STRONG LOCKEAN THESIS (SLT). There is a threshold $\frac{1}{2} < s < 1$ such that all rational $\langle \mathcal{B}, p \rangle$ satisfy

$$\mathcal{B}(A) \text{ iff } p(A) \geq s.$$

WEAK LOCKEAN THESIS (WLT). For every rational $\langle \mathcal{B}, p \rangle$ there is a threshold $\frac{1}{2} < s < 1$ such that

$$\mathcal{B}(A) \text{ iff } p(A) \geq s.$$

Most discussions of the Lockean thesis have in mind the strong thesis. More recent work, especially Leitgeb (2017), adopts the weaker thesis. The strong thesis leaves the correct threshold unspecified. Of course for every $\frac{1}{2} < s < 1$, we can formulate a specific thesis SLT^s in virtue of which the strong thesis is true. For example, $\text{SLT}^{.51}$ is a very permissive version of the thesis, whereas $\text{SLT}^{.95}$ and $\text{SLT}^{.99}$ are more stringent. It is also possible to further specify the weak thesis. For example, Leitgeb (2017) believes that the contextually-determined threshold should be equal to the degree of belief assigned to the strongest proposition that is fully believed. In light of Deductive Cogency, that corresponds to the orthographically ungainly $\text{WLT}^{p(\cap \mathcal{B})}$.

The strong Lockean thesis gives rise to the well-known *Lottery paradox*, due originally to Kyburg (1961, 1997). The lesson of the Lottery is that the strong thesis is in tension with Deductive Cogency. Suppose that s is the universally correct Lockean threshold. Now think of a fair lottery with N tickets, where N is chosen large enough that $1 - (1/N) \geq s$. Since the lottery is fair, it seems permissible to fully believe that *some* ticket is the winner. It also seems reasonable to assign degree of belief $1/N$ to each proposition of the form "The i^{th} ticket is the winner." According to the Lockean thesis, such an agent ought to fully believe that the first ticket is a loser, the second ticket is a loser, the third is a loser, etc. Since cogency requires belief to be closed under conjunction, she ought to believe that all the tickets are losers. But now she violates cogency by believing both that every ticket is a loser and that some ticket is a winner. Since s was arbitrary, we have shown that no matter how high we set the threshold, there is some Lottery for which an agent must either violate the Lockean thesis or violate Deductive Cogency. According to Kyburg, what the paradox teaches is that we should give up on Deductive Cogency: full belief should not necessarily be closed under conjunction. Many others take the lesson of the Lottery to be that the strong Lockean thesis is untenable.

Several authors attempt to revise the strong Lockean thesis by placing restrictions on when a high degree of belief warrants full belief. Broadly speaking, they propose that a high degree of belief is sufficient to warrant full belief unless some defeating condition holds. For example, Pollock (1995) proposes that, although a degree of belief in P above some threshold is a *prima facie* reason for belief, that reason is defeated whenever P is a member of an inconsistent set of propositions each of which is also believed to a degree exceeding the threshold. Ryan (1996) proposes that a high degree of belief is sufficient for full belief unless the proposition is a member of a set of propositions such that each member has a degree of belief exceeding the threshold, but the probability of their conjunction is below the threshold. Douven (2002) says that it is sufficient except when the proposition is a member of a *probabilistically self-undermining* set. A set $\mathcal{S}$ is probabilistically self undermining iff for all $A \in \mathcal{S}, p(A) > s$ and $p(A \mid B) \leq s$, where $B = \bigcap(\mathcal{S} \setminus \{A\})$. It is clear that any of these proposals would prohibit full belief that a particular lottery ticket will lose.

These proposals are all vitiated by the following sort of example due to Korb (1992). Let A be any proposition with a degree of belief above threshold but short of certainty. Let L_i be the proposition that the i^{th} lottery ticket (of a large lottery with N tickets) will lose. Consider the set $\mathcal{S} = \{\neg A \cup L_i \mid 1 \leq i \leq N\}$. Each member of $\mathcal{S}$ is above threshold, since L_i is above threshold. Furthermore, the set $\mathcal{S} \cup \{A\}$ meets all three defeating conditions. Therefore, these proposals prohibit full belief in any proposition with degree of belief short of certainty. Douven and Williamson (2006) generalize this sort of example to trivialize an entire class of similar formal proposals.

Buchak (2014) argues that what partial beliefs count as full beliefs cannot merely be a matter of the degree of partial belief, but must also depend on the type of evidence it is based on. According to Buchak, this means there can be no merely formal answer to the question: what conditions on partial belief are necessary and sufficient for full belief? The following example, of a type going back to Thomson (1986), illustrates the point. Your parked car was hit by a bus in the middle of the night. The bus could belong either to the blue bus company or the red bus company. Consider the following two scenarios.

SCENARIO 1. You know that the blue company operates 90% of the buses in the area, and the red bus company operates only 10%. You have degree of belief 0.9 that a blue bus is to blame.

SCENARIO 2. The red and blue companies operate an equal number of buses. A 90% reliable eyewitness testifies that a blue bus hit your car. You have degree of belief 0.9 that a blue bus is to blame.

Buchak (2014) argues that it is rational to have full belief that a blue bus is to blame in the second scenario, but not in the first. You have only statistical evidence in the first scenario, whereas in the second, a causal chain of events connects your belief to the accident (see also Thomson, 1986, Nelkin, 2000, and Schauer, 2003). These intuitions, Buchak observes, are reflected in our legal practice: purely statistical evidence is not sufficient to convict. If you find Buchak's point convincing, you will be unsatisfied with most of the proposed accounts for how full and partial belief ought to correspond (Staffel, 2016).

Despite difficulties with buses and lotteries, the dynamics of qualitative belief under the strong thesis are independently interesting to investigate. For example, van Eijck and Renne (2014) axiomatize the logic of belief for a Lockean with threshold $\frac{1}{2}$. Makinson and Hawthorne (2015) investigate which principles of non-monotonic logic are validated by Lockean agents. Before turning to proposed solutions to the Lottery paradox, we make some observations about qualitative Lockean revision, inspired largely by Shear and Fitelson (2018).

It is a theorem of the probability calculus that $p(H \mid E) \leq P(E \to H)$. So if H is assigned a high degree of belief given E, the material conditional $E \to H$ must have been assigned a degree of belief at least as high *ex ante*. It is easy to see that as a probabilistic analogue of the principle of Conditionalization from non-monotonic logic or, equivalently, the AGM Inclusion principle. That observation has the following consequence: any belief that the Lockean comes to have after conditioning, she could have arrived at by adding the evidence to her prior beliefs and closing under logical consequence. Therefore Lockean updating satisfies the AGM principle of Inclusion. Furthermore, it follows immediately from definitions that Lockean update satisfies Success and Extensionality.

Theorem 13 *Suppose that* $s \in (\frac{1}{2}, 1)$. *Let* $\mathcal{B}_E = \{A : p(A \mid E) \geq s\}$ *for all* $E \in \mathcal{F}_p^+$. *Then* $\mathcal{B}_{\mathcal{F}_p^+}$ *satisfies Inclusion, Success, and Extensionality.*

In Section 2.2.1, we argued that Inclusion and Preservation capture the spirit of AGM revision. If Lockean revision also satisfied Preservation, we would have a clean sweep of the AGM principles, with the exception of Deductive Cogency.

However, that cannot hold in general. It is possible to construct examples where $p(\neg E) < s$, $p(H) \geq s$, and yet $p(H \mid E) < s$. For Lockean agents this means that it is possible to lose a belief, even when revising on a proposition that is not disbelieved.

Recall the example of Alice, Bob, and the Ford from Section 2.1.1. Let $W = \{a, b, c\}$ corresponding to the worlds in which Alice owns the Ford, Bob owns the Ford, and no one in the office owns the Ford. Suppose the probability function

$$p(a) = \frac{6}{10},$$
$$p(b) = \frac{3}{10},$$
$$p(c) = \frac{1}{10},$$

captures my partial beliefs. For Lockean thresholds in the interval $(.75, .9]$, my full beliefs are exhausted by $\mathcal{B} = \{\{a, b\}, W\}$. Now suppose I were to learn that Alice does not own the Ford. That is consistent with all beliefs in $\mathcal{B}$, but since $p(\{a, b\} \mid \{b, c\}) = \frac{3}{4}$, it follows by the Lockean thesis that $\{a, b\} \notin \mathcal{B}_{\{b,c\}}$. So Lockeanism does not in general validate Preservation. The good news, at least for those sympathetic to Pollock's critique of non-monotonic logic, is that the Lockean thesis allows for undercutting defeat of previous beliefs.

However, Shear and Fitelson (2018) also have some good news for fans of AGM and the Lockean thesis. Two quantities are in the *golden ratio* ϕ if their ratio is the same as the ratio of their sum to the larger of the two quantities, i.e. for $a > b > 0$, if $\frac{a+b}{a} = \frac{a}{b}$ then $\frac{a}{b} = \phi$. The golden ratio is an irrational number approximately equal to 1.618. Its inverse ϕ^{-1} is approximately .618. Shear and Fitelson prove the following intriguing result.

Theorem 14 *Suppose that* $s \in (\frac{1}{2}, \phi^{-1}]$. *Let* $\mathcal{B}_E = \{A : p(A \mid E) \geq s\}$ *for all* $E \in \mathcal{F}_p^+$. *Let*

$$\mathcal{D} = \{E \subseteq W : E \in \mathcal{F}_p^+ \textit{ and } \mathcal{B}_E \textit{ is deductively cogent}\}.$$

Then $\mathcal{B}_{\mathcal{D}}$ *satisfies the six basic AGM postulates.*

That shows that for relatively low thresholds, Lockean updating satisfies all the AGM postulates—at least when we restrict to deductively cogent belief sets.

Why has the golden ratio turned up here? That is relatively simple to explain. The AGM Preservation postulate can be factored into the following two principles.

$$\text{If } \neg E \notin Cn(\mathcal{B}) \text{ and } E \in Cn(\mathcal{B}) \text{ then } \mathcal{B} \subseteq \mathcal{B}_E. \quad \text{(Cautious Monotony)}$$

$$\text{If } \neg E \notin Cn(\mathcal{B}) \text{ and } E \notin Cn(\mathcal{B}) \text{ then } \mathcal{B} \subseteq \mathcal{B}_E. \quad \text{(Preservation B)}$$

We have discussed Cautious Monotony in Section 2.1.1. It is widely accepted as a *sine qua non* of rational non-monotonic reasoning. Surprisingly, there is no Lockean threshold that satisfies Cautious Monotony in general.[8] However, if $p(H \mid E) < s$ it must be that $p(H \cap E) < s \cdot P(E) \leq s$, from which it follows that any violation of Cautious Monotony must be a violation of deductive closure. Moreover, Lockean updating with a threshold in $(\frac{1}{2}, \phi^{-1}]$ satisfies Preservation B. That follows immediately from the fact that for $s \in (\frac{1}{2}, \phi^{-1}]$, if $p(E) < s$ and $p(H \mid E) < s$, then $P(H \to \neg E) \geq s$. The proof of that fact hinges on a neat fact about the golden ratio: if $s > 0$, then $s \leq \phi^{-1}$ iff $s^2 \leq 1 - s$.[9]

5.3 *The Stability Theory of Belief*

For many, sacrificing Deductive Cogency is simply too high a price to pay for a bridge principle, even one so simple and intuitive as the strong Lockean thesis. That occasions a search for bridge principles that can be reconciled with Deductive Cogency. One proposal, due to Leitgeb (2013, 2014, 2015, 2017) and Arló-Costa and Pedersen (2012), holds that rational full belief corresponds to a stably high degree of belief, i.e. a degree of belief that remains high even after conditioning on new information. Leitgeb calls this view the *Humean thesis*, due to Hume's conception of belief as an idea of superior vivacity, but also of superior steadiness.[10] Leitgeb (2017) formalizes Hume's definition, articulating the following version of the thesis:

Humean Thesis (HT). For all rational pairs $\langle \mathcal{B}, p \rangle$ there is $s \geq 1/2$ such that

$$\mathcal{B}(A) \text{ iff } \neg B \notin \mathcal{B} \text{ implies } p(A \mid B) > s.$$

8 See Lemma 1 in Shear and Fitelson (2018).

9 Suppose that $s \in (\frac{1}{2}, \phi^{-1}]$ and $P(E) < s$ and $P(H \mid E) < s$. Then, $P(E)P(H \mid E) = P(H \cap E) < s^2 \leq 1 - s$, and therefore $1 - P(H \cap E) = P(H \to \neg E) \geq s$.

10 See Loeb (2002, 2010) for a detailed development of the stability theme in Hume's conception of belief.

In other words: every full belief must have stably high conditional degree of belief, at least when conditioning on propositions which are not currently disbelieved. Since full belief occurs on both sides of the biconditional, it is evident that this is not a proposed *reduction* of full belief to partial belief, but rather a constraint that every rational agent must satisfy. The Humean thesis leaves the precise threshold s unspecified. Of course for every $\frac{1}{2} < s < 1$, we can formulate a specific thesis HT^s in virtue of which the thesis is true. For example, $\mathrm{HT}^{.5}$ requires that every fully believed proposition remains more likely than its negation when conditioning on propositions not currently disbelieved.

Some form of stability is widely considered to be a necessary condition for *knowledge*. Socrates propounds such a view in the *Meno*. Paxson and Lehrer (1969) champion such a view in the epistemology literature post-Gettier. However, stability is not usually mooted as a condition of *belief*. Raidl and Skovgaard-Olsen (2017) claim that Leitgeb's stability condition is more appropriate in an analysis of knowledge and too stringent a condition on belief. A defender of the Humean thesis might say that every *rational* belief is possibly an instance of knowledge. Since knowledge is necessarily stable, unstable beliefs are *ipso facto* not known.

Leitgeb demonstrates the following relationships between the Humean thesis, Deductive Cogency, and the weak Lockean thesis.

Theorem 15 *Suppose that $\langle \mathcal{B}, p \rangle$ satisfy* HT *and $\varnothing \notin \mathcal{B}$. Then, $\mathcal{B}$ is deductively cogent and $\langle \mathcal{B}, p \rangle$ satisfy* $\mathrm{WLT}^{p(\cap\mathcal{B})}$.

So if an agent satisfies the Humean thesis and does not "fully" believe the contradictory proposition, her qualitative beliefs are deductively cogent and furthermore, she satisfies the weak Lockean thesis, where the threshold is set by the degree of belief assigned to $\cap\mathcal{B}$, the logically strongest proposition she believes. Leitgeb also proves the following partial converse.

Theorem 16 *Suppose that $\mathcal{B}$ is deductive cogent and $\langle \mathcal{B}, p \rangle$ satisfy* $\mathrm{WLT}^{p(\cap\mathcal{B})}$. *Then, $\langle \mathcal{B}, p \rangle$ satisfy* $\mathrm{HT}^{\frac{1}{2}}$ *and $\varnothing \notin \mathcal{B}$.*

Together, these two theorems say that the Humean thesis (with threshold $\frac{1}{2}$) is equivalent to Deductive Cogency and the weak Lockean thesis (with threshold $p(\cap\mathcal{B})$). Since it is always possible to satisfy $\mathrm{HT}^{\frac{1}{2}}$, Leitgeb gives us an ingenious way to reconcile Deductive Cogency with a version of the Lockean thesis.

Recall the example of the lottery. Let $W = \{w_1, w_2, \ldots, w_N\}$, where w_i is the world in which the i^{th} ticket is the winner. No matter how many tickets are in the lottery, a Humean agent cannot believe any ticket will lose. Suppose for a contradiction that she believes $W \setminus \{w_1\}$, the proposition that the first ticket will lose. Now suppose she learns $\{w_1, w_2\}$, that all but the first and second ticket will lose. This is compatible with her initial belief, but her updated degree of belief that the first ticket will lose must be $\frac{1}{2}$. That contradicts the Humean thesis. So she cannot believe that any ticket will lose. In this Lottery situation the agent cannot fully believe any non-trivial proposition. This example also shows how sensitive the Humean proposal is to the fine-graining of possibilities. If we coarsen W into the set of possibilities $W = \{w_1, w_2\}$, where w_1 is the world in which the first ticket is the winner and w_2 is "the" world in which some other ticket is the winner, the agent can believe that the first ticket will lose without running afoul of the Humean thesis.

Perhaps Buchak (2014) is right and no agent should have beliefs in lottery propositions—these beliefs would necessarily be formed on the basis of purely statistical evidence. Kelly and Lin (forthcoming) give another scenario in which Humean agents seem radically skeptical, but in situations which are evidentially unproblematic. Suppose the luckless Job goes in for a physical. On the basis of a thorough examination, the doctor forms the following dire opinion of his health: her degree of belief that Job will survive exactly n months is $\frac{1}{2^n}$. Therefore, her degree of belief that Job will not survive the year is $\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{12}} > .999$. Shockingly, the Humean thesis prevents the doctor from forming *any* nontrivial beliefs. Let $\leq n$ be the proposition that Job survives at most n months and let $\geq n$ be the proposition that he survives at least n months. Let B be the strongest proposition that the doctor believes. Suppose for a contradiction that B entails some least upper bound for the number of Job's remaining months, i.e for some n, B entails $\leq n$ and does not entail $\leq n'$ for any $n' < n$. By construction, $p(B| \geq n) = p(n)/p(\geq n) = \frac{1}{2}$ for all n. But since $\geq n$ is compatible with B, the Humean thesis requires that $p(B| \geq n) > \frac{1}{2}$. Contradiction.

The example of the doctor suggests that the price of Humeanism is a rather extreme form of skepticism: in many situations a Humean agent will have no non-trivial full beliefs at all. That criticism is developed extensively in Rott (2017) and Douven and Rott (2018). The doctor also illustrates how the Humean proposal allows arbitrarily small perturbations of partial beliefs to be reflected as huge differences in full beliefs. Suppose the doctor is slightly more confident that Job will not survive a month, i.e. her survival probabilities decrease as $\frac{1}{2} + \epsilon, \frac{1}{4}, \frac{1}{8} - \epsilon, \frac{1}{16}, \frac{1}{32}, \ldots$. Now the doctor can believe that Job will be dead in two months without running afoul of the Humean thesis.

So far we have inquired only into the synchronic content of the Humean proposal. What sort of principles of qualitative belief update does it underwrite? Leitgeb demonstrates an intimate relationship between the AGM revision principles and the Humean thesis: every agent that satisfies the AGM principles, as well as a weak version of the Lockean thesis, must also satisfy the Humean thesis. So if you think that AGM theory is the correct theory of rational qualitative belief update (and you believe that a high degree of partial belief is a *necessary* condition of full belief) you must also accept the Humean thesis.

To present Leitgeb's result we have to introduce a few technical concepts. Say that a proposition A is p-stable^r iff for all $B \in \mathcal{F}_p^+$ such that $A \cap B \neq \varnothing$, $p(A \mid B) > r$. An immediate consequence of this definition is that if A is p-stable^r and A is consistent with $E \in \mathcal{F}_p^+$, then $A \cap E$ is p_E-stable^r. Let

$$\mathcal{S}_p^r = \{A : A \text{ is } p\text{-stable}^r\}.$$

Leitgeb proves that for $r \geq 1/2$, the set $\mathcal{S}_p^r$ is a system of spheres in the sense of Section 2.2.3. That is: there is some least element B of $\mathcal{S}_p^r$ such that all other elements constitute a nested, well-ordered sphere system centered on B. Recall that $\mathcal{S}_p^r(E)$ is defined to be $D \cap E$, where D is the closest sphere to B compatible with E. By the previous observation, $\mathcal{S}_p^r(E)$ is p_E-stable^r.

Leitgeb proves the following.

Theorem 17 *The following are equivalent.*

1. $\mathcal{B}_{\mathcal{F}_p^+}$ *satisfies all AGM postulates and for all* $E \in \mathcal{F}_p^+$, $A \in \mathcal{B}_E$ *only if* $p(A \mid E) > r$.
2. $\cap \mathcal{B}_E = \mathcal{S}_p^r(E) \in \mathcal{S}_{p_E}^r$.

We know from the result of Section 2.2.3 that for any AGM belief revision operation, there is a corresponding system of Grove spheres. Leitgeb has proven that any agent that validates the AGM postulates and the high-probability requirement can be modeled by the system of spheres generated by the p-stable^r propositions. For such an agent, all pairs $\langle \mathcal{B}_E, p_E \rangle$ satisfy the Humean thesis with threshold r. So any agent that violates the Humean thesis must either fail to satisfy the AGM postulates, or the high-probability requirement. Note that the converse is not true: it is not the case that that if all pairs $\langle \mathcal{B}_E, p_E \rangle$ satisfy the Humean thesis, then $\mathcal{B}_{\mathcal{F}_p^+}$ must satisfy the AGM postulates. To prove this, suppose that $\langle \mathcal{B}, p \rangle$ satisfy the Humean thesis and $\cap \mathcal{B} \subset E$ for some $E \in \mathcal{F}_p^+$. If we let $\mathcal{B}_E = \{E\}$, then $\langle \mathcal{B}_E, p_E \rangle$ satisfy the Humean thesis. However, such an agent patently violates Rational and even Cautious Monotony.

5.4 *The Tracking Theory*

Lin and Kelly (2012) propose that qualitative belief update ought to *track* partial belief update. On their picture, partial and full beliefs are maintained and updated by parallel cognitive systems. The first system, governed by the probabilistic norms of Bayesian coherence and conditioning, is precise, slow, and cognitively expensive. That system is engaged for important deliberations requiring a lot of precision and occurring without much time pressure e.g. retirement planning. The second, which in some way maintains and updates full beliefs, is quicker and less cognitively burdensome.[11] That system is engaged in ordinary planning: grocery shopping, or selecting a restaurant for a department event. What keeps these two parallel systems in sync with each other?

Lin and Kelly study *acceptance rules* that specify a mechanism for transitioning gracefully into the qualitative and out of the probabilistic system. An acceptance rule α maps every partial belief state p to a unique qualitative belief state $\alpha(p)$ with which it coheres. For example, the strong Lockean thesis determines an acceptance rule once we specify a threshold. The Humean thesis, on the other hand, underdetermines an acceptance rule, merely imposing constraints on acceptable pairs $\langle \mathcal{B}, p \rangle$. An agent's qualitative updates *track* her probabilistic updates iff

$$\alpha(p)_E = \alpha(p_E),$$

whenever $p(E) > 0$. In other words: acceptance followed by qualitative revision yields the same belief state as probabilistic revision followed by acceptance.

Here is a way to understand the tracking requirement. Suppose that, although an agent maintains a latent probabilistic belief state, most of her cognitive life is spent reasoning with and updating qualitative beliefs. A typical day will go by without having to engage the probabilistic system at all. Suppose Monday is a typical day. Let $\langle \alpha(p), p \rangle$ be the belief state she wakes up with on Monday: her full and partial beliefs are in harmony. Let E be the total information she acquired since waking up. Since qualitative beliefs are updated on the fly, she goes to sleep with the qualitative belief state $\alpha(p)_E$. Overnight, her probabilistic system does the difficult work of Bayesian conditioning and computes the partial belief state p_E, just in case she runs into any sophisticated decision problems on Tuesday. Before waking, she transitions out of her probabilistic system p_E and into the qualitative belief state $\alpha(p_E)$. If she fails the tracking requirement, she may wake up on Tuesday morning with a qualitative belief state that is drastically different from the one she went to sleep with on Monday night. If she tracks, then she will notice no difference at all. For such an agent, no mechanism (other than memory) is required to bring her full and partial beliefs back into harmony on Tuesday morning. Supposing that we *enter* the probabilistic system by conditioning our previous partial belief state p on all new information E, and *exit* by accepting $\alpha(p_E)$, tracking ensures that transitioning in and out of the probabilistic system does not induce

11 For an objection to the two systems view, see Staffel (2018).

any drastic changes in qualitative beliefs. An agent that tracks will notice no difference at all. An agent that does not track may find her full and partial beliefs perpetually falling out of sync, requiring many expensive acceptance operations to bring them back into harmony.

Tracking may be a desirable property, but are there any architectures that exhibit it? Lin and Kelly (2012) answer this question affirmatively. Since Bayesian conditioning is taken for granted, Lin and Kelly must specify two things: a qualitative revision operation and an acceptance rule that jointly track conditioning. We turn now to the details of their proposal. As usual, let W be a set of worlds. A *question* $\mathcal{Q}$ is a partition of W into a countable collection of mutually exhaustive propositions $H_1, H_2, \ldots$, which are the complete *answers* to $\mathcal{Q}$. The partial belief function p is defined over the algebra of propositions $\mathcal{A}$ generated by $\mathcal{Q}$.

First we specify an acceptance rule. Lin and Kelly propose the *odds threshold rule*. The degree of belief function p is used to determine a plausibility order by setting

$$H_i \prec_p H_j \qquad \text{if and only if} \qquad \frac{p(H_i)}{p(H_j)} > t,$$

where t is a constant greater than 1 and $p(H_i), p(H_j) > 0$. This determines an acceptance rule by setting $\alpha(p) = \mathcal{B}_{\prec_p}$. Since the odds threshold rule determines a plausibility order $\prec_p$ and any plausibility order $\prec$ gives rise to a deductively cogent belief state $\mathcal{B}_{\prec}$, the Lottery paradox is avoided. In other words: the bridge principle that any rational $\langle \mathcal{B}, p \rangle$ are related by $\mathcal{B} = \alpha(p)$ ensures that $\mathcal{B}$ is deductively cogent. Furthermore, the odds threshold rule allows non-trivial qualitative beliefs in situations where the stability theory precludes them. Recall the case of the doctor. Consider the odds threshold $2^{10} - 1$. Given this threshold, the hypothesis that Job will survive exactly 1 month is strictly more plausible than the proposition that he will survive at least n months for any $n \geq 10$. This threshold yields the full belief that Job will survive at most 10 months. However, in the case of the Lottery the odds threshold rule precludes any non-trivial beliefs.[12] See Rott (2017) and Douven and Rott (2018) for an extensive comparison of the relative likelihood of forming non-trivial qualitative beliefs on the odds-threshold and stability proposals.

It remains to specify the qualitative revision operation. Lin and Kelly adopt an operation proposed by Shoham (1987). Let $\prec$ be a well-founded, strict partial order over the answers to $\mathcal{Q}$.[13] This is interpreted as a *plausibility ordering*, where $H_i \prec H_j$ means that H_i is strictly *more* plausible than

12 The content-dependent threshold rule proposed by Kelly and Lin (forthcoming) may allow non-trivial beliefs in the Lottery situation.

13 A strict partial order is *well-founded* iff every subset of the order has a least element. This is closely related to the stopperedness property discussed in Section 2.1.2.

H_j. Every plausibility order $\prec$ gives rise to a belief state $\mathcal{B}_\prec$ by letting $\neg H_i \in \mathcal{B}_\prec$ iff there is some H_j strictly more plausible than H_i and closing under logical consequence. In other words, $\cap\mathcal{B}_\prec$ is the disjunction of the minimal elements in the plausibility order. The plausibility order $\prec$ is updated on evidence E by setting every answer incompatible with E to be strictly less plausible than every answer compatible with E, and otherwise leaving the order unchanged. Let $\prec_E$ denote the result of this update operation. We use the updated plausibility order to define a belief revision rule by setting $\mathcal{B}_E = \mathcal{B}_{\prec_E}$. Then, for all $E, F \subseteq W$, B_E is deductively cogent and satisfies:

$$\cap\mathcal{B}_E \subseteq E; \qquad \text{(Success)}$$

$$\cap\mathcal{B} \cap E \subseteq \cap\mathcal{B}_E; \qquad \text{(Inclusion)}$$

$$\text{if } \cap\mathcal{B} \subseteq E \text{ then } \cap\mathcal{B}_E \subseteq \cap\mathcal{B}. \qquad \text{(Cautious monotony)}$$

However, it does not necessarily satisfy Preservation. To see this suppose that $\mathcal{Q} = \{H_1, H_2, H_3\}$ and $H_1 \prec H_2$ but H_3 is not ordered with H_1 or H_2. Then $\cap\mathcal{B} = H_1 \cup H_3$. However $\cap\mathcal{B}_{\neg H_1} = H_2 \cup H_3 \nsubseteq \cap\mathcal{B}$ even though $\cap\mathcal{B} \cap \neg H_1 \neq \varnothing$.

Lin and Kelly prove that Shoham revision and odds-threshold based acceptance jointly track conditioning.

Theorem 18 *Let $\prec$ equal $\prec_p$ and let $\mathcal{B}_E = \mathcal{B}_{\prec_E}$. Then $\mathcal{B}_{\wp(W)}$ satisfies Deductive Cogency, Success, Cautious Monotony, and Inclusion. Furthermore, $\mathcal{B}_E = \alpha(p)_E = \alpha(p_E)$ for all $E \in \mathcal{F}_p^+$.*

In other words: odds-threshold acceptance followed by Shoham revision yields the same belief state as Bayesian conditioning followed by odds-threshold acceptance.[14] Although the original plausibility ordering $\prec_p$ is built from the probability function p, subsequent qualitative update proceeds without consulting the (conditioned) probabilities. That shows that there are at least some architectures that effortlessly keep the probabilistic and qualitative reasoning systems in harmony.

Fans of AGM will regret that Shoham revision does not satisfy AGM Preservation (Rational Monotony). Lin and Kelly (2012) prove that no "sensible" acceptance rule that tracks conditioning can satisfy Inclusion and Preservation. According to Lin and Kelly, sensible acceptable rules are *non-skeptical*, *non-opinionated*, *consistent*, and *corner-monotonic*. An acceptance rule is *non-skeptical* iff for every answer H_i to $\mathcal{Q}$ there is a non-negligible

14 Kelly and Lin (forthcoming) recommend a modification of the odds-threshold rule proposed in Lin and Kelly (2012).

set of probability functions p such that $H_i \in \alpha(p)$.[15] An acceptance rule is *non-opinionated* iff there is a non-negligible set of probability functions p where judgement is suspended, i.e. where $\cap\alpha(p) = W$. An acceptance rule is *consistent* iff for all p, $\alpha(p)$ is deductively cogent. The intuition behind corner-monotony is that if H_i is accepted at p, then H_i should still be accepted if H_i is made more probable. More precisely, an acceptance rule is *corner-monotone* iff $H_i \in \alpha(p)$ implies that $H_i \in \alpha(p')$ for all p' such that

$$p' = p(\cdot \mid H_i) \cdot q + p(\cdot \mid \neg H_i) \cdot (1 - q),$$

and $q > p(H_i)$. Lin and Kelly (2012) prove the following "no-go" theorem for AGM revision.

Theorem 19 *Suppose that $\mathcal{B}_E = \alpha(p_E)$ for $E \in \mathcal{F}_p^+$. Then $\mathcal{B}_{\mathcal{F}_p^+}$ satisfies Inclusion and Preservation only if α is not sensible.*

5.5 *Decision-Theoretic Accounts*

All of the bridge principles we have seen so far have the following in common: whether an agent's full and partial beliefs cohere is a matter of the full and partial beliefs *alone*. It is not necessary to mention preferences or utilities in order to evaluate a belief state. There is another tradition, originating in Hempel (1962) and receiving classical expression in Levi (1967), that assimilates the problem of "deciding" what to believe to a Bayesian decision-theoretic model. Crucially, these authors are not committed to a picture on which agents literally decide what to believe—rather they claim that an agent's beliefs are subject to the same kind of normative evaluation as their practical decision-making. Contemporary contributions to this tradition include Easwaran (2015), Pettigrew (2016c), and Dorst (2017). Presented here is a somewhat simplified version of Levi's (1967) account taking propositions, rather than sentences, as the objects of belief.

As usual, let W be a set of possible worlds. The agent is taken to be interested in answering a *question* $\mathcal{Q}$, which is a partition of W into a finite collection of mutually exhaustive answers $\{H_1, H_2, \ldots H_n\}$. Levi calls situations of this sort "efforts to replace agnosticism by true belief," echoing themes in Peirce (1877).

> Doubt is an uneasy and dissatisfied state from which we struggle to free ourselves and pass into the state of belief; while the

15 A set of probability functions is *non-negligible* iff it contains an open set in the topology generated by the metric

$$||p - q|| = \sqrt{\sum_{H_i \in \mathcal{Q}} (p(H_i) - q(H_i))^2}.$$

> latter is a calm and satisfactory state which we do not wish to avoid, or to change to a belief in anything else. On the contrary, we cling tenaciously, not merely to believing, but to believing just what we do believe.

The agent's partial beliefs are represented by a probability function p that is defined, at a minimum, over the algebra $\mathcal{A}$ generated by the question. Levi recommends the following procedure to determine which propositions are fully believed: disjoin all those elements of $\mathcal{Q}$ that have *maximal* expected epistemic utility and then close under deductive consequence. The *expected epistemic utility* of a hypothesis $H \in \mathcal{A}$ is defined as:

$$E(H) := p(H) \cdot U(H) + p(\neg H) \cdot u(H),$$

where $U(H)$ is the epistemic utility of accepting H when it is true, and $u(H)$ is the utility of accepting H when it is false. How are $u(H), U(H)$ to be determined? Levi is guided by the following principles.

1. True answers have greater epistemic utility than false answers.

2. True answers that afford a high degree of relief from agnosticism have greater epistemic utility than true answers that afford a low degree of relief from agnosticism.

3. False answers that afford a high degree of relief from agnosticism have greater epistemic utility than false answers that afford a low degree of relief from agnosticism.

It is easy to object to these principles. The first principle establishes a lexicographic preference for true beliefs. It is conceivable that, contra this principle, an informative false belief that is approximately true should have greater epistemic utility than an uninformative true belief. The first principle precludes trading content against truthlikeness. It is also conceivable that, contra the third principle, one would prefer to be wrong, but not too opinionated, than wrong and opinionated. The only unexceptionable principle seems to be the second.

To measure the degree of relief from agnosticism, a probability function $m(\cdot)$ is defined over the elements of $\mathcal{A}$. Crucially, $m(\cdot)$ does not measure a degree of belief, but degree of *uninformativeness*. The degree of relief from agnosticism afforded by $H \in \mathcal{A}$, also referred to as the *amount of content* in H, is defined to be the complement of uninformativeness: $cont(H_i) = m(\neg H_i)$. Levi argues that all the elements of $\mathcal{Q}$ ought to be assigned the same amount of content, i.e. $m(H_i) = \frac{1}{n}$ and therefore $cont(H_i) = \frac{n-1}{n}$ for each $H_i \in \mathcal{Q}$. The set of epistemic utility functions that Levi recommends satisfy the following conditions:

$$\begin{aligned} U(H) = &\quad 1 - q \cdot cont(\neg H), \\ u(H) = &\quad -q \cdot cont(\neg H), \end{aligned}$$

where $0 < q < 1$. All such utility functions are guaranteed to satisfy Levi's three principles. The parameter q is interpreted as a "degree of caution," representing the premium placed on truth as opposed to relief from agnosticism. When $q = 1$ the epistemic utility of suspending judgement, $U(W)$, is equal to zero. This is the situation in which the premium placed on relief from doubt is the maximum. Levi proves that expected epistemic utility $E(H)$ is maximal iff $p(H) > q \cdot cont(\neg H)$. Therefore, Levi's ultimate recommendation is that the agent believe all deductive consequences of

$$\bigcap\{\neg H_i \in \mathcal{Q} : p(\neg H_i) > 1 - q \cdot cont(\neg H_i)\}.$$

From this formulation it is possible to see Levi's proposal as a question-dependent version of the Lockean thesis where the appropriate threshold is a function of content. However, Levi takes pains to make sure that the result of this operation is deductively cogent and therefore avoids Lottery-type paradoxes.

Contemporary contributions to the decision-theoretic tradition proceed differently from Levi. Most recent work does not take epistemic utility to be primarily a function of content. Most of these proposals do not refer to a question in context. Many proposals, such as Easwaran (2015) and Dorst (2017), are equivalent to a version of the Lockean thesis, where the threshold is determined by the utility the agent assigns to true and false beliefs. Since these are essentially Lockean proposals, they are subject to Lottery-style paradoxes.

REFERENCES

Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *The journal of symbolic logic*, *50*(2), 510–530.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica*, *21*(4), 503–546.

Arló-Costa, H. (1999). Qualitative and probabilistic models of full belief. In S. R. Buss, P. Hájek, & P. Pudlák (Eds.), *Proceedings of logic colloquium* (Vol. 98, pp. 25–43).

Arló-Costa, H. & Pedersen, A. P. (2012). Belief and probability: A general theory of probability cores. *International Journal of Approximate Reasoning*, *53*(3), 293–315.

Armendt, B. (1980). Is there a Dutch book argument for probability kinematics? *Philosophy of Science*, *47*, 583–588.

Briggs, R. A. (2017). Normative theories of rational choice: Expected utility. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University.

Buchak, L. (2014). Belief, credence, and norms. *Philosophical Studies*, *169*(2), 285–311.
Cohen, L. J. (1977). *The probable and the provable*. Oxford: Clarendon Press.
Cohen, L. J. (1980). Some historical remarks on the Baconian conception of probability. *Journal of the History of Ideas*, *41*(2), 219–231.
Cohen, L. J. (1992). *An essay on belief and acceptance*. Clarendon Press: Oxford.
de Finetti, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. In *Annales de l'institut henri poincaré* (Vol. 7, *1*, pp. 1–68).
de Finetti, B. (1970). *Theory of probability*. New York: Wiley.
de Finetti, B. (1972). *Probability, induction and statistics*. New York: Wiley.
Dempster, A. P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society (Series B, Methodological)*, *30*(2), 205–247.
Dorst, K. (2017). Lockeans maximize expected accuracy. *Mind*, *128*(509), 175–211.
Douven, I. (2002). A new solution to the paradoxes of rational acceptability. *British Journal for the Philosophy of Science*, *53*, 391–410.
Douven, I. & Rott, H. (2018). From probabilities to categorical beliefs: Going beyond toy models. *Journal of Logic and Computation*, *28*(6), 1099–1124.
Douven, I. & Williamson, T. (2006). Generalizing the lottery paradox. *British Journal for the Philosophy of Science*, *57*, 755–779.
Doyle, J. (1979). A truth maintenance system. *Artificial intelligence*, *12*(3), 231–272.
Doyle, J. (1992). Reason maintenance and belief revision: Foundations vs. coherence theories. In P. Gärdenfors (Ed.), *Belief revision* (pp. 29–52). Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.
Earman, J. (1992). *Bayes or bust?: A critical examination of bayesian confirmation theory*. MIT Press.
Easwaran, K. (2011a). Bayesianism I: Introduction and arguments in favor. *Philosophy Compass*, *6*(5), 312–320.
Easwaran, K. (2011b). Bayesianism II: Applications and criticisms. *Philosophy Compass*, *6*(5), 321–332.
Easwaran, K. (2015). Dr. Truthlove or: How I learned to stop worrying and love Bayesian probabilities. *Nous*, *50*(4), 816–853.
Easwaran, K. (2019). Conditional probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.
Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, *75*(4), 643–669.
Eriksson, L. & Hájek, A. (2007). What are degrees of belief? *Studia Logica*, *86*(2), 183–213.
Fine, T. (1973). *Theories of probability: An examination of foundations*. Elsevier.

Foley, R. (1993). *Working without a net: A study of egocentric epistemology*. Oxford University Press.

Frankish, K. (2009). Partial belief and flat-out belief. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of belief* (pp. 73–79). Synthese Library. Springer.

Gabbay, D. M. (1985). Theoretical foundations for non-monotonic reasoning in expert systems. In *Logics and models of concurrent systems* (pp. 439–457). Springer.

Gärdenfors, P. (1986). The dynamics of belief: Contractions and revisions of probability functions. *Topoi*, *5*(1), 29–37.

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

Gärdenfors, P. (1992). Belief revision: An introduction. In P. Gärdenfors (Ed.), *Belief revision* (pp. 1–29). Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.

Gärdenfors, P. & Makinson, D. (1988). Revisions of knowledge systems using epistemic entrenchment. In *Proceedings of the 2nd conference on theoretical aspects of reasoning about knowledge* (pp. 83–95). Morgan Kaufmann Publishers Inc.

Genin, K. (2017). How inductive is Bayesian conditioning? *Manuscript*. Retrieved from https://kgenin.github.io/papers/conditioning%5C_long.pdf

Gettier, E. L. (1963). Is justified true belief knowledge? *analysis*, *23*(6), 121–123.

Grove, A. (1988). Two modellings for theory change. *Journal of philosophical logic*, *17*(2), 157–170.

Gyenis, Z., Hofer-Szabó, G., & Rédei, M. (2017). Conditioning using conditional expectations: The Borel-Kolmogorov paradox. *Synthese*, *194*(7), 2595–2630.

Hacking, I. (1975). *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge University Press.

Hájek, A. (2003). What conditional probability could not be. *Synthese*, *137*(3), 272–323.

Hájek, A. (2012). Interpretations of probability. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University.

Hájek, A. & Lin, H. (2017). A tale of two epistemologies? *Res Philosophica*, *94*(2), 207–232.

Hansson, S. O. (1999). *A textbook of belief dynamics: Theory change and database updating*. Kluwer Academic Publishers.

Hansson, S. O. (2017). Logic of belief revision. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2017). Metaphysics Research Lab, Stanford University.

Harman, G. (1986). *Change in view: Principles of reasoning*. MIT Press.

Hempel, C. G. (1962). Deductive-nomological vs. statistical explanation. In H. Fiegl & G. Maxwell (Eds.), *Vol 3.* (pp. 98–169). Minnesota Studies in the Philosophy of Science. University of Minnesota Press.

Horgan, T. (2017). Troubles for bayesian formal epistemology. *Res Philosophica*, *94*(2), 233–255.

Horty, J. F. (2012). *Reasons as defaults*. Oxford University Press.

Howson, C. & Urbach, P. (2006). *Scientific reasoning: The bayesian approach*. Open Court Publishing.

Huber, F. (manuscript). *Belief and counterfactuals. a study in means-end philosophy.* Oxford University Press.

Huber, F. (2013a). Belief revision I: The AGM theory. *Philosophy Compass*, *8*(7), 604–612.

Huber, F. (2013b). Belief revision II: Ranking theory. *Philosophy Compass*, *8*(7), 613–621.

Huber, F. (2016). Formal representations of belief. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2016). Metaphysics Research Lab, Stanford University.

Huber, F. (2019). Ranking theory. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Jeffrey, R. C. (1970). Dracula meets Wolfman: Acceptance vs. Partial Belief. In M. Swain (Ed.), *Induction, acceptance, and rational belief*. Synthese Library. D. Reidl.

Jeffrey, R. C. (1983). *The logic of decision* (2nd). University of Chicago Press.

Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science*, *65*, 575–603.

Kaplan, M. (1996). *Decision theory as philosophy*. Cambridge University Press.

Kelly, K. T. (1996). *The logic of reliable inquiry*. Oxford University Press.

Kelly, K. T. & Lin, H. (forthcoming). Beliefs, probabilities, and their coherent correspondence. In I. Douven (Ed.), *Lotteries, knowledge and rational belief: Essays on the lottery paradox*. Cambridge University Press.

Kemeny, J. G. (1955). Fair bets and inductive probabilities. *Journal of Symbolic Logic*, *20*, 263–273.

Keynes, J. M. (1921). A treatise on probability. vol. 8 of collected writings (1973 ed.) London: Macmillan.

Kolmogorov, A. N. (1950). *Foundations of the theory of probability*. New York: Chelsea.

Konek, J. (2019). Comparative probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Koopman, B. O. (1940). The axioms and algebra of intuitive probability. *The Annals of Mathematics*, *41*(2), 269–292.

Korb, K. B. (1992). The collapse of collective defeat: Lessons from the lottery paradox. In D. Hull, M. Forbes, & K. Okruhlik (Eds.), *Psa: Proceedings of the biennial meeting of the philosophy of science association* (Vol. 1, pp. 230–236).

Kraus, S., Lehmann, D., & Magidor, M. (1990). Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial intelligence*, *44*(1-2), 167–207.

Kvanvig, J. L. (2016). Intellectual humility: Lessons from the preface paradox. *Res Philosophica*, *93*(3), 509–532.

Kyburg, H. E. (1961). *Probability and the logic of rational belief*. Wesleyan University Press.

Kyburg, H. E. (1997). The rule of adjunction and reasonable inference. *The Journal of Philosophy*, *94*(3), 109–125.

Lehrer, K. (1965). Knowledge, truth and evidence. *Analysis*, *25*(5), 168–175.

Leitgeb, H. (2013). Reducing belief simpliciter to degrees of belief. *Annals of Pure and Applied Logic*, *164*(12), 1338–1389.

Leitgeb, H. (2014). The stability theory of belief. *The Philosophical Review*, *123*(2), 131–171.

Leitgeb, H. (2015). I—the Humean thesis on belief. In *Aristotelian society supplementary volume* (Vol. 89, *1*, pp. 143–185). Wiley Online Library.

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

Leitgeb, H. & Pettigrew, R. (2010). An objective justification of Bayesianism II: The consequences of minimizing inaccuracy. *Philosophy of Science*, *77*, 236–272.

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

Levi, I. (1977). Subjunctives, dispositions and chances. *Synthese*, *34*(4), 423–455.

Levi, I. (1991). *The fixation of belief and its undoing: Changing beliefs through inquiry*. Cambridge University Press.

Lewis, D. (1979). Attitudes de dicto and de se. *The philosophical review*, *88*(4), 513–543.

Lewis, D. (1999). Why conditionalize? In *Papers in metaphysics and epistemology* (pp. 403–407). Cambridge University Press.

Liao, S.-y. (2012). What are centered worlds? *The Philosophical Quarterly*, *62*(247), 294–316.

Lin, H. (2013). Foundations of everyday practical reasoning. *Journal of Philosophical Logic*, *42*(6), 831–862.

Lin, H. (2019). Belief revision theory. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Lin, H. & Kelly, K. T. (2012). Propositional reasoning that tracks probabilistic reasoning. *Journal of philosophical logic*, *41*(6), 957–981.
Loeb, L. E. (2002). *Stability and justification in hume's treatise*. Oxford University Press on Demand.
Loeb, L. E. (2010). *Reflection and the stability of belief: Essays on Descartes, Hume, and Reid*. Oxford University Press on Demand.
Mahtani, A. (2019). Imprecise probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.
Makinson, D. (1965). The paradox of the preface. *Analysis*, *25*(6), 205–207.
Makinson, D. (1994). General patterns in nonmonotonic reasoning. In D. M. Gabbay, C. Hogger, & J. Robinson (Eds.), *Handbook of logic in artificial intelligence and logic programming (vol. 3): Nonmonotonic reasoning and uncertain reasoning* (pp. 35–111). Oxford University Press.
Makinson, D. & Hawthorne, J. (2015). Lossy inference rules and their bounds: A brief review. In *The road to universal logic* (pp. 385–407). Springer.
Moon, A. (2017). Beliefs do not come in degrees. *Canadian Journal of Philosophy*, *47*(6), 760–778.
Moss, S. (2018). *Probabilistic knowledge*. Oxford University Press.
Nelkin, D. K. (2000). The lottery paradox, knowledge, and rationality. *The Philosophical Review*, *109*(3), 373–409.
Paxson, T., Jr. & Lehrer, K. (1969). Knowledge: Undefeated justified true belief. *The Journal of Philosophy*, *66*(8), 225–237.
Peirce, C. S. (1877). The fixation of belief. In N. Houser & C. Kloesel (Eds.), *The essential Pierce: Selected philosophical writings. vol. i.* Indiana University Press.
Perry, J. (1979). The problem of the essential indexical. *Noûs*, 3–21.
Pettigrew, R. (2016a). *Accuracy and the laws of credence*. Oxford University Press.
Pettigrew, R. (2016b). Epistemic utility arguments for probabilism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University.
Pettigrew, R. (2016c). Jamesian epistemology formalised: An explication of 'The Will to Believe'. *Episteme*, *13*(3), 253–268.
Pollock, J. L. (1987). Defeasible reasoning. *Cognitive Science*, *11*(4), 481–518.
Pollock, J. L. (1995). *Cognitive carpentry: A blueprint for how to build a person*. MIT Press.
Pollock, J. L. (2006). *Thinking about acting: Logical foundations for rational decision making*. Oxford University Press.
Popper, K. (1955). Two autonomous systems for the calculus of probabilities. *British Journal for the Philosophy of Science*, *6*(3-4), 51–57.
Popper, K. & Miller, D. (1983). A proof of the impossibility of inductive probability. *Nature*, *302*(5910), 687.

Quine, W. V. O. (1990). *Pursuit of truth*. Harvard University Press.
Quine, W. V. O. & Ullian, J. S. (1970). *The web of belief*. Random House.
Raidl, E. (forthcoming). Completeness for counter-doxa conditionals—using ranking semantics. *The Review of Symbolic Logic*.
Raidl, E. & Skovgaard-Olsen, N. (2017). Bridging ranking theory and the stability theory of belief. *Journal of Philosophical Logic*, *46*(6), 577–609.
Ramsey, F. P. (1931). Truth and probability. In R. Braithwaite (Ed.), *The foundations of mathematics and other logical essays* (pp. 156–199). Routledge.
Renyi, A. (1955). On a new axiomatic system for probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, *6*, 285–335.
Resnik, M. D. (1987). *Choices: An introduction to decision theory*. University of Minnesota Press.
Roorda, J. (1995). Revenge of Wolfman: A probabilistic explication of full belief. *unpublished*. Retrieved from https://www.princeton.edu/~bayesway/pu/Wolfman.pdf
Ross, D. (1930). *The right and the good*. Oxford University Press.
Rott, H. (2003). Basic entrenchment. *Studia Logica*, *73*(2), 257–280.
Rott, H. (2017). Stability and scepticism in the modelling of doxastic states: Probabilities and plain beliefs. *Minds and Machines*, *27*(1), 167–197.
Ryan, S. (1996). The epistemic virtues of consistency. *Synthese*, *109*, 121–141.
Savage, L. J. (1954). *The foundations of statistics*. Wiley publications in statistics. Wiley.
Schauer, F. (2003). *Profiles, probabilities and stereotypes*. Belknap Press.
Schurz, G. (2011). Abductive belief revision in science. In *Belief revision meets philosophy of science* (pp. 77–104). Springer.
Schwitzgebel, E. (2015). Belief. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*.
Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press.
Shear, T. & Fitelson, B. (2018). Two approaches to belief revision. *Erkenntnis*.
Shoham, Y. (1987). A semantical approach to nonmonotonic logics. In *Readings in nonmonotonic reasoning* (pp. 227–250). Morgan Kaufmann Publishers Inc.
Skyrms, B. (2009). Diachronic coherence and probability kinematics. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of belief* (pp. 73–79). Synthese Library. Springer.
Spohn, W. (1988). Ordinal conditional functions: A dynamic theory of epistemic states. In W. L. Harper & B. Skyrms (Eds.), *Causation in decision, belief change and statistics: Proceedings of the irvine conference*

on probability and causation (pp. 105–134). The University of Western Ontario Series in Philosophy of Science. Kluwer.

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

Spohn, W. (2017). Knightian uncertainty meets ranking theory. *Homo Oeconomicus*, *34*(4), 293–311.

Spohn, W. (2019). Defeasible normative reasoning. *Synthese*.

Staffel, J. (2013). Can there be reasoning with degrees of belief? *Synthese*, *190*(16), 3535–3551.

Staffel, J. (2016). Beliefs, buses and lotteries: Why rational belief can't be stably high credence. *Philosophical Studies*, *173*(7), 1721–1734.

Staffel, J. (2018). How do beliefs simplify reasoning? *Noûs*.

Stalnaker, R. (1981). Indexical belief. *Synthese*, *49*(1), 129–151.

Stalnaker, R. (1984). *Inquiry*. MIT Press.

Stalnaker, R. (1994). What is a nonmonotonic consequence relation? *Fundamenta Informaticae*, *21*(1, 2), 7–21.

Strasser, C. & Antonelli, G. A. (2018). Non-monotonic logic. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University.

Teller, P. (1973). Conditionalization and observation. *Synthese*, *26*, 218–258.

Thoma, J. (2019). Decision theory. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Thomson, J. J. (1986). Liability and individualized evidence. *Law and Contemporary Prolems*, *49*(3), 199–219.

Ullman-Margalit, E. (1983). On presumption. *The Journal of Philosophy*, *80*(3), 143–163.

van Fraassen, B. C. (1995). Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical logic*, *24*(4), 349–377.

van Eijck, J. & Renne, B. (2014). Belief as willingness to bet. *CoRR*, *abs/1412.5090*. arXiv: 1412.5090. Retrieved from http://arxiv.org/abs/1412.5090

von Neumann, J. & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Weirich, P. (2004). Belief and acceptance. In I. Niiniluoto, M. Sintonen, & J. Woleński (Eds.), *Handbook of epistemology* (pp. 499–520). Dordrecht: Kluwer Academic Publishers.

Weisberg, J. (2011). Varieties of bayesianism. In D. M. Gabbay, S. Hartmann, & J. Woods (Eds.), *Volume 10: Inductive logic* (pp. 477–553). Handbook of the History of Logic. Elsevier.