# 1

# Introduction

We begin in §1.1 with an informal and intuitive introduction to objective Bayesianism as the view that an agent's degrees of belief should be probabilities, should respect constraints imposed by empirical evidence, but should otherwise equivocate between basic propositions. This is followed in §1.2 by an outline of the positive claims of the book and in §1.3 by an overview of the key challenges and objections faced by objective Bayesianism, and pointers to where in the book each of these objections are met. Then §1.4 sets out some of the key assumptions of the book concerning an agent's evidence, her language, and rationality.

## 1.1 Objective Bayesianism outlined

*Objective Bayesian epistemology* (also known simply as *objective Bayesianism*) is a theory about how strongly one should believe propositions that are open to speculation. It is a normative theory, concerning the strength with which one *ought* to believe certain propositions, rather than a descriptive theory about one's actual degrees of belief. This book presents a version of objective Bayesianism that is characterized by the following tenets. The strengths of an agent's beliefs should behave like probabilities: they should be representable by real numbers in the unit interval and one should believe a disjunction of mutually exclusive propositions to the extent of the sum of the degrees of belief of the disjuncts. Moreover, these degrees of belief should be shaped by empirical evidence: for example, they should be calibrated with known frequencies. Where empirical evidence does not fully determine degrees of belief, they should equivocate. Thus, the version of objective Bayesianism presented here can be summed up as follows: *degrees of belief should be probabilistic, calibrated with evidence, and otherwise equivocal*.

Consider, for example, a consultant who needs to determine a prognosis for a patient who has breast cancer. The consultant knows only that the patient has been treated for breast cancer, and that in somewhere between 10% and 40% of such cases the cancer recurs. Objective Bayesianism then maintains the following. The degree to which the agent should believe that the patient's cancer will recur ($R$) is representable by a probability $P(R) \in [0, 1]$, where $P(R) + P(\neg R) = 1$. This degree of belief should be constrained by the frequency evidence, that is, $P(R) \in [0.1, 0.4]$. Finally, the consultant should equivocate as far as possible between $R$ and $\neg R$, and therefore should set $P(R) = 0.4$ and $P(\neg R) = 0.6$.

The aim of this book is to provide a detailed development of objective Bayesianism and to defend the resulting theory against a broad range of criticisms.

### 1.2 Objective Bayesian theory

This book is as much about theory-building as about meeting objections to objective Bayesianism. Here we set the scope of objective Bayesian theory as developed in the book.

The orthodox view of Bayesian epistemology goes something like this. All Bayesian epistemologists hold that rational degrees of belief are probabilities, consistent with total available evidence, and updated in the light of new evidence by Bayesian conditionalization. Strict subjectivists (e.g. Bruno de Finetti) hold that initial or *prior* degrees of belief are largely a question of personal choice. Empirically based subjectivists (e.g. Howson and Urbach 1989*a*) hold that prior degrees of belief should not only be consistent with total evidence, but should also be calibrated with physical probabilities to the extent that they are known. Objectivists (e.g. Edwin Jaynes) hold that prior degrees of belief are fully determined by the evidence.

This book preserves the distinction between the three kinds of Bayesian epistemologies, but departs from orthodoxy inasmuch as it develops a version of objective Bayesianism that is not updated by conditionalization, and where probabilities are not fully determined by evidence—they are not always fully determined, and where they are, they are determined by more than evidence alone.

According to the version of objective Bayesian epistemology developed here, an agent's degrees of belief should depend on her evidence, her language, and various other contextual factors. As explained in §1.4, the agent's evidence includes everything she takes for granted in her current operating context. Her language not only allows her to express the propositions of interest to her, but also conceptualizes the salient features of her world and delimits the range of possible circumstances that she can consider. We restrict our attention to three kinds of formal languages in this book: the language of propositional logic (§2.3), the language of predicate logic (§5.1), and the language of the mathematical theory of probability (§9.1).

In this book, objective Bayesianism is characterized in terms of three norms. The Probability norm says that the strengths of an agent's beliefs should be representable by a probability function defined over the sentences of the agent's language. The Calibration norm says that this probability function should fit with the agent's evidence. The Equivocation norm says that, to the extent that evidence leaves open a range of evidentially compatible probability functions, the agent's degrees of belief should be representable by some such function that equivocates to a sufficient extent between the basic propositions expressible in her language. The different versions of Bayesianism are distinguishable according to their stance regarding these norms: while objective Bayesianism advocates all three norms, empirically based subjectivists accept Probability and Calibration but reject Equivocation, and strict subjectivists only endorse the Probability norm. The book explicates these norms in a more precise way and shows that they do indeed yield strong constraints on degrees of belief. Chapter 3 discusses the motivation behind the three norms, arguing that any attempt at rigorously

proving a norm will be futile in the sense that such a proof is unlikely to convince a detractor. However, it is argued that the norms do admit compelling pragmatic justifications. So, while the norms cannot be proven and need to be suitably qualified, all three norms are genuine constraints on rational belief and not merely conventions (§3.5).

The book focuses on explicating the Equivocation norm through the versions of Edwin Jaynes' Maximum Entropy Principle (see §2.3), although §3.4 leaves open the possibility of other explications. This yields a natural mechanism for updating degrees of belief which differs from the subjective Bayesian method of updating, Bayesian conditionalization (Chapter 4). It also makes possible a relatively efficient way of calculating objective Bayesian probabilities, the machinery of objective Bayesian nets developed in Chapter 6.

Although in its infancy in comparison with other interpretations of probability (§2.2), objective Bayesian theory already admits many applications, for example to decision support systems (§6.3), to logic (Chapter 7), to epistemology (Chapter 8), to statistics (§10.2), to scientific reasoning (§10.3), and to metaphysics (§10.4). It is hoped that the book will encourage others to further develop objective Bayesian theory and its applications.

### 1.3 Criticisms of objective Bayesianism

Objective Bayesianism has been challenged on a number of different fronts.

It has been accused of being poorly motivated. For example, Festa (1993, §7.2) criticizes the Maximum Entropy Principle on the grounds that it is not clear how evidence constrains probability. It is hoped that by being explicit about the three norms of objective Bayesianism, and being explicit about the extent and limitations of the motivation behind each of these norms (Chapter 3), many of these worries can be allayed.

Objective Bayesianism has been accused of poorly handling qualitative evidence. For example, Pearl (1988, p. 463) maintains that the Maximum Entropy Principle fails to adequately handle causal evidence. This criticism is addressed in §3.3 as well as in Williamson (2005*a*). No doubt there is more to do in this direction, but as things stand various kinds of qualitative evidence—including evidence of causal, logical, semantic, and ontological influence—can now be handled.

Objective Bayesianism has been accused of yielding updates that can disagree with Bayesian conditionalization (see, e.g. Friedman and Shimony 1971). It is shown in Chapter 4 that while this is indeed the case, it reflects negatively on Bayesian conditionalization rather than on objective Bayesian updating.

Objective Bayesianism has been accused of suffering from a failure to learn from experience (see, e.g. Dias and Shimony 1981, §4; Paris 1994, p. 178). It is shown in §4.3 that this objection rests on a mistaken understanding of objective Bayesian conditional probabilities.

Objective Bayesianism has been accused of being computationally intractable. For instance, Pearl (1988, p. 463) dismisses the Maximum Entropy

Principle on these grounds. Chapter 6 shows that Pearl's own Bayesian network formalism can achieve a significant dimension reduction in the entropy maximization task, and can render objective Bayesianism computationally tractable.

Objective Bayesianism has been accused of being susceptible to the paradoxes of the Principle of Indifference: there may be more than one partition of sentences over which one might equivocate, each partition giving different conclusions (see, e.g. Seidenfeld 1986, §3). In §9.1, it is acknowledged that there may in certain cases be more than one potential equivocator—a function that is maximally equivocal over the basic propositions expressible in the language—but at worst this leads to subjectivity rather than paradox.

Relatedly, objective Bayesianism has been accused of being language-dependent, that is, of yielding probabilities that depend on an agent's language as well as her evidence (see, e.g. Seidenfeld 1986, §3). This is indeed the case, but only to be expected, since an agent's language says something about the world in which she lives—it embodies implicit evidence (§9.2).

One might object to objective Bayesianism that it is not objective enough: there are several situations in which the three norms fail to determine a unique probability function, and, in any case, objective Bayesian probability is relative to evidence and language. In fact, §9.3 shows that subjectivity in the former sense is by no means pernicious and that relativity to evidence and language can be eliminated if need be.

No doubt there are some criticisms of objective Bayesianism in the literature that do not fall under one of the aforementioned broad objections. But those set out above have been the most influential criticisms and are the principal reasons invoked by subjectivists for preferring subjective over objective Bayesian epistemology. By showing that these objections can be substantially met, the book aims to level the playing field.

The book contains no sustained attack on subjectivism, many of which have been conducted elsewhere, but in defending objective Bayesianism several considerations that cast doubt on subjectivism will emerge.

## 1.4 Evidence, language, and rationality

In this section, we set out and discuss some of the presuppositions that pervade the rest of the book.

### 1.4.1 *Evidence*

Objective Bayesianism posits a link between an agent's evidence and her rational degrees of belief: degrees of belief should be constrained by the evidence and otherwise maximally equivocal. *Evidence* is understood here in a broad sense to include *everything the agent takes for granted* in her current operating context—for example, observations, theory, background knowledge, and assumptions. The terms 'total evidence' or 'epistemic background' or 'data' are sometimes used to refer to what is called 'evidence' in this book.

Note that an agent's evidence is relative to her current operating context. A medical consultant, for example, is rational to take for granted much of her education and medical training as well as uncontroversial studies in the medical literature and observations of the patient's symptoms, given her purpose of treating a patient. On the other hand, the purposes of a philosophical sceptic preclude her from taking much for granted at all. Therefore, if the medical consultant by day studies philosophy by night, her evidence base may change radically.

The agent's evidence is denoted by the symbol $\mathcal{E}$, and the probability function that represents the degrees of belief that the agent should adopt is denoted by $P_{\mathcal{E}}$. There is no need to assume that the evidence $\mathcal{E}$ is a set of propositions, or that the agent is always aware of what she grants—we can leave these questions open. Indeed, since objective Bayesianism only concerns the *link between* $\mathcal{E}$ and $P_{\mathcal{E}}$, there is no need to say very much about $\mathcal{E}$ at all, other than to explain how it constrains $P_{\mathcal{E}}$. Note, in particular, that an agent may or may not be rational to take $\mathcal{E}$ for granted, but objective Bayesianism says nothing about this question. Indeed, it is important to emphasize that the theory of objective Bayesianism does not aim to answer all questions about epistemic rationality: it is but one piece in the jigsaw puzzle of normativity.

It is also important to emphasize that an agent's evidence includes more than her background knowledge. Evidence (understood as everything the agent takes for granted in her current context) differs from knowledge in three respects. First, an agent may take something for granted when she ought not: for example, she may take some testimony for granted although she is aware that it is probably idle gossip. Such evidence does not qualify as knowledge because it lacks the required justification, or because it is arrived at through an unreliable or otherwise inappropriate process.[1] Second, an agent may take something for granted—and be justified in so doing—yet this something fails to qualify as knowledge because it is false: for example, she may take some testimony from a very reliable witness for granted, but it just turns out to be a rare occasion in which the witness was mistaken. Third, it may be the case that the agent does not take an item of her knowledge for granted: perhaps she should have taken it for granted but didn't, or perhaps she was right not to take it for granted because it was irrelevant in her operating context or because her operating context placed more stringent demands than does knowledge (e.g. that a reliable witness makes a true claim may be sufficient for that claim to be counted as knowledge, but not sufficient for it to be counted as evidence presentable to a court of law).

Not everyone agrees that evidence (taken as that which grounds belief) and knowledge differ. For example, Timothy Williamson maintains that knowledge and evidence are coextensive (Williamson 2000, chapter 9). Williamson

[1] Of course it is controversial as to how knowledge itself should be understood. However, accounts of knowledge standardly maintain that truth is necessary but insufficient for knowledge, so that there are at least two ways in which a knowledge claim might be false—it may lack truth or may lack whatever else is required for it to count as knowledge. It is this 'whatever else' that is intended here.

suggests two arguments for the claim that evidence must be true. First, 'if one's evidence included falsehoods, it would rule out some truths, by being inconsistent with them. One's evidence may make some truths improbable, but it should not exclude any outright' (Williamson 2000, p. 201). In response, one might agree that it is desirable that all one's evidence be true, but point out that mere desirability does not make it so. A possible second argument concerns a different *relational* notion of evidence—evidence *for* a hypothesis:

**1**: 'If $e$ is evidence for $h$, then $e$ is true. . . . For example, that the ground is wet is evidence that it rained last night only if the ground is wet' (Williamson 2000, p. 201).

A link is posited between the two notions of evidence:

**2**: $e$ is evidence for $h$ for an agent if and only if it is in the agent's body of evidence, and $e$ raises the probability of $h$ (Williamson 2000, p. 187).

Suppose that one can make a case that

**3**: if $e$ is in the agent's body of evidence, then there is some $h$ whose probability $e$ raises.

Then one could conclude that $e$ being in the agent's body of evidence implies that it is true. In response to this second argument one can challenge any one of its premisses. Consider the claim, common in discussions of computational complexity in computer science, that $P \neq NP$. This statement has not been proven and it is universally agreed that it might be false, yet the overwhelming majority of computer scientists grant it, and grant it for many good reasons. Now consider premise 1. It is commonly agreed that $P \neq NP$ is evidence that the travelling salesman problem is computationally intractable. Now $P \neq NP$ may be false. Should it be found that $P = NP$, then arguably the evidential claim—the claim that $P \neq NP$ is evidence for the travelling salesman problem being intractable—would be considered irrelevant rather than false. It would be what Williamson calls 'conversationally inappropriate' to utter such a claim, and in response to such a claim one would point to $P \neq NP$ being false, rather than the claim as a whole being false. To rebut premise 2, note that $P \neq NP$ does not raise the probability of the travelling salesman problem being intractable, on any standard account of probability including Williamson's own and that of this book, since any standard account gives $P \neq NP$ probability 0 or 1. Indeed, $P \neq NP$ raises the probability of no proposition so premise 3 must be false. In sum, neither of these two arguments forces the conclusion that evidence is knowledge.

Discussions of Bayesian epistemology in the literature often take background knowledge to be the sole constrainer of degrees of belief. In fact, though, one's degrees of belief are—and should be—guided by what one grants rather than what one knows. If knowledge ought to guide an agent when she is consciously deciding how strongly to believe a proposition that is open to speculation, then, since *ought* implies *can*, the agent should be able to identify exactly what it is she

knows. But of course an agent cannot be expected to do this, because while she may think she knows something, she may be mistaken about its justification or its truth. On the other hand, the agent can identify what she takes for granted (at least with some training and time to reflect) and so can use her evidence to guide the strengths of her beliefs. Of course *can* does not imply *ought*: that evidence (taken as whatever the agent grants) can guide belief does not imply that it should. In particular, while one may accept that knowledge is too strong a condition on the grounds for belief, one might also suspect that evidence is too strong a condition. Might one not base belief on something weaker than what is *taken for granted*? On other beliefs perhaps? Thus, believing a proposition strongly (without taking it for granted) may be grounds for believing its negation weakly. In response to this worry it suffices to point out that while the proposition itself is not granted and hence does not qualify as evidence, *believing the proposition strongly* is granted and so does qualify as evidence. Hence, the *grounds* for believing the negation of the proposition weakly are granted rather than merely believed. In general, the grounds for one's beliefs need to be taken for granted if they are to ground those beliefs at all: if $X$ warrants epistemic state $E$ and one adopts state $E$ on the basis of $X$, then one had better grant $X$. Consequently, it is evidence (construed as that which one grants), not something stronger or weaker, that should guide the strengths of one's beliefs.

That it is evidence not knowledge that should ground belief is apparent from examples in the legal setting where much effort has been directed at understanding belief formation. In a court of law the jurors should apportion their beliefs not according to their knowledge but according to the evidence under consideration: some of this evidence may in fact be false and hence not count as knowledge; for it to guide the jurors' beliefs, it is sufficient that they grant it. On the other hand, for technical reasons the judge may instruct that something the jurors have heard and known be discounted as evidence and should not influence their beliefs. It is not just belief formation in the law that behaves like this: in most contexts it is appropriate that our observations guide our beliefs, even those that turn out to be erroneous; in many contexts in engineering and science it is appropriate to grant Newtonian mechanics, even though it is known to be strictly false; an observation of a planet by the naked eye may yield knowledge of the planet, yet in the context of astronomical theorizing more precise and sure observation may be required in order for that knowledge to be granted, and hence guide the ascription of belief.

### 1.4.2 *Language*

While we can leave open the possibility that elements of the agent's evidence be non-propositional, it will be assumed that the objects of the agent's degrees of belief are propositions. It is the agent's *language*, which we will usually denote by $\mathcal{L}$, that picks out these propositions. Just as evidence may vary according to the agent's operating context, so too her language may vary with context. Thus, a structural engineer who moonlights as a baker may be thought of as having two

languages whose technical terms intersect little. We need not assume anything about the relation between degrees of beliefs on different languages of the same agent where the context varies; if the languages intersect, one may want degrees of belief on the different languages to be consistent on their intersection, or some weaker paraconsistency may be sufficient, but there is no need to decide such issues here. On the other hand, the methods of this book will apply to the case in which an agent's language *in a single context* changes over time.

It will be helpful for the most part to consider formal languages: we shall assume that if the language can express propositions $\theta, \varphi$ then it can express $\neg\theta$ (not $\theta$), $\theta \wedge \varphi$ ($\theta$ and $\varphi$), $\theta \vee \varphi$ ($\theta$ or $\varphi$), $\theta \rightarrow \varphi$ (not $\theta$, or $\varphi$), and $\theta \leftrightarrow \varphi$ (both or neither of $\theta$ and $\varphi$). We shall at various points of this book consider languages that have finitely many, or denumerably many, or uncountably many elementary (i.e. non-logically complex) sentences, and we shall in some cases exploit the fine-grained structure of sentences by considering first-order logical languages with predicate, constant, variable, and quantifier symbols.

To say that an agent has language $\mathcal{L}$ and evidence $\mathcal{E}$ is not to say that there is no relation between evidence and language. While we do not presume that an agent's evidence $\mathcal{E}$ can be expressed in her language $\mathcal{L}$, we do not rule this out. Moreover, in Chapter 9 we shall see that the agent's language itself constitutes a kind of implicit evidence, conveying information about the world in which she operates. Thus, in principle the agent's language as well as her other evidence $\mathcal{E}$ can constrain her degrees of belief. Where we need to emphasize this dependence of degrees of belief on language as well as evidence, we shall denote the agent's belief function by $P^{\mathcal{L}}_{\mathcal{E}}$.

### 1.4.3 *Rationality*

This book focuses on *rational* degree of belief. Here 'rational' may simply be taken to mean *fit-for-purpose*. (These terms are clearly not inter-substitutable, though, because they have different domains of application: 'rational' is a normative notion and is applied to things over which we have some control—such as the strengths of our beliefs in propositions that are open to speculation—while 'fit-for-purpose' is descriptive and more widely applicable, for example, to artifacts, such as tables and chairs.)

One's degrees of belief are rational inasmuch as they are fit for the uses to which they are put. These uses are principally truth-orientated—for example, making predictions that turn out true, assessing which diagnosis is the right one, and determining the correct explanation. This book will only be concerned with degrees of belief in the context of these truth-orientated uses—these uses broadly require that degrees of belief fit reality as far as possible. Particular individuals may have ulterior motives when deciding what to believe. For instance, if one wants beliefs to lead to happiness, self-delusion may be the appropriate strategy in so far as it is possible to carry out. Interesting as these ancillary uses are, they will not be pursued here.

The type of rationality under consideration here is sometimes called *theoretical rationality* or *epistemic rationality* to distinguish it from *practical rationality* or *pragmatic rationality*. There are some grounds for drawing a distinction here (Harman 2004), but it must be drawn with caution. In particular, Gilbert Harman goes perhaps too far by arguing that theoretical reasoning is not characterizable as practical reasoning with a theoretical goal, on the grounds that practical reasoning involves arbitrary decisions (in deciding how to achieve a result there may be several equally promising means to that end, requiring an arbitrary choice of which means to adopt), while theoretical reasoning does not (Harman 2007, §4). We shall see in this book that theoretical reasoning *can* involve arbitrary choice: there may be several equally rational degrees of belief from which the agent should choose one arbitrarily. Furthermore, we shall see that even though the *uses* of degrees of belief may be truth-orientated and hence there is a sense in which reasoning with rational degrees of belief qualifies as theoretical reasoning, this kind of reasoning can also be characterized as reasoning with a *practical goal* and hence there is also a sense in which it is practical reasoning. This is because for degrees of belief to be deemed rational, they need to avoid the possibility of certain loss in both the short term (§3.2) and the long term (§3.3), and to minimize worst-case expected loss (§3.4), all of which are practical ends. In general, while many attempt to drive a wedge between belief and action (see, e.g. Kyburg Jr 1970), they are but two sides of the same coin—beliefs are the epistemic determinants of action—and one can straightforwardly draw inferences from one to the other (§§3.1, 3.5). This truism helps to muddy the water in distinguishing theoretical from practical rationality, a phenomenon known as *pragmatic encroachment* (see, e.g. Fantl and McGrath 2007). Pragmatic encroachment need not detain us here. By embracing the pragmatic aspects of belief but distinguishing its truth-orientated *uses* from ulterior motives, we have enough of a distinction to focus on truth-orientated rationality in this book.

In sum, rational degree of belief is relative to evidence and language. Degrees of belief are rational if they are fit-for-purpose. Plausibly, degrees of belief are rational if and only if they are determined in the right way from evidence and language, and if in the first place the agent is rational to grant that evidence and adopt that language (in so far as the agent has any choice about her evidence and language). Objective Bayesian epistemology concerns only the former question: the link between evidence and language on the one hand, and rational degree of belief on the other; it does not concern choice of evidence or choice of language. Objective Bayesianism is a theory which holds that the degrees of belief which best fit their purpose are those that are probabilistic, calibrated with evidence, and otherwise equivocal. In Chapter 3, we examine some of the reasons for invoking these principles, but first we shall take a look at some of the key developments on the path to present-day objective Bayesianism.