UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

# ON THE METHODOLOGICAL AND NORMATIVE FOUNDATIONS OF PROBABILISM

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

PATRICK G. EPLEY
Norman, Oklahoma
2015

ON THE METHODOLOGICAL AND NORMATIVE FOUNDATIONS OF PROBABILISM

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PHILOSOPHY

BY

______________________________
Dr. James Hawthorne, Chair

______________________________
Dr. Stephen Ellis

______________________________
Dr. Martin Montminy

______________________________
Dr. Wayne Riggs

______________________________
Dr. Robert Terry

Dedicated to the memory of

Isabelle O. Horvath

## Acknowledgements

I had a lot of help and support writing this dissertation and finishing this degree. So, I owe many debts of gratitude—too many to list here. However, I would be remiss if I did not make the following acknowledgements.

My two greatest debts of gratitude are to my spouse, Kelly Epley, and my dissertation director, Jim Hawthorne. Kelly put up with hours and hours of excited talk about probability and degrees of belief, and she supported me through times of great stress, all the while pursuing her own doctoral studies. I hope to repay this care and support in kind as she finishes her own dissertation.

Jim has been an outstanding advisor and mentor to me since he first took me on as a PhD student. He has been extremely generous with his time, energy, and intellect, and he has given me invaluable advice, support, and feedback through every stage of the degree and numerous drafts of this dissertation.

I also owe many thanks to my dissertation committee members, Stephen Ellis, Martin Montminy, Wayne Riggs, and Robert Terry for generously offering me their time, encouragement, and feedback.

This project began to take shape while I was a visiting student at Rutgers University in fall 2012. Thanks to Branden Fitelson for supervising this visit.

I would also like thank the many members of the OU Philosophy Department for providing me with an unparalleled environment in which to pursue a PhD in philosophy. Thanks especially to Shelley Konieczny and Gabe Serrano.

Finally, I would also like to thank my family and friends for supporting me through this long, stressful process. In particular, I would like to thank my mother, Veronica Epley, for editing the penultimate draft of this dissertation. It brings me great joy to know that she has read every page!

## Table of Contents

## List of Tables

## List of Figures

## Abstract

This dissertation is an elaboration and defense of probabilism, the view that belief comes in various degrees of strength, and that the probability calculus provides coherence norms for these degrees of belief. Probabilism faces several well-known objections. For example, critics object that probabilism's numerical representation of degrees of belief is too precise, and that its coherence norms are too demanding for real human agents to follow. While probabilists have developed several plausible responses to these objections, the compatibility among these responses is unclear. On this basis, I argue that probabilists must articulate unified methodological and normative foundations for their view, and I sketch the foundations of a probabilist modeling framework, the Comparative Confidence Framework (CCF). CCF characterizes probabilism primarily as an account of ideal degree of belief coherence. CCF provides a set of fundamentally qualitative and comparative—rather than quantitative—evaluative ideals for degree of belief coherence. By providing qualitative, comparative, evaluative coherence norms for degrees of belief, CCF avoids the aforementioned objections: that probabilism's formal representation of degrees of belief is too precise, and that its norms are too demanding. CCF is a first step in the development of unified foundations for a wider subjectivist Bayesian theory of doxastic and pragmatic rationality.

# 1 Introduction

## 1.1 Précis of the Project

This dissertation is about probabilism. In particular, as the title indicates, it's about probabilism's methodological and normative foundations. Put very broadly, probabilism is a philosophical application of the probability calculus—the mathematical theory of probability—to characterize norms for apportioning our confidence in the truth of claims. Philosophical applications of probability have become very popular recently. Philosophers have been using probabilistic tools in the philosophy of science and decision theory for most of the last century. But, recently, metaphysicians, philosophers of language, ethicists, and "mainstream" epistemologists have also begun to employ probabilistic tools with increasing regularity.[1] So, reflection on the philosophical foundations underlying this tool is clearly in order.

Probabilism is one of the core components of subjective Bayesianism,[2] a highly fruitful and influential research program in epistemology, philosophy of science,

---

[1]See Titelbaum (2013, p. 3) for a more detailed list of some recent philosophical applications of probability.

[2]I write of "subjective Bayesianism" to refer to the version of Bayesianism that uses probabilities primarily to represent—or normatively model—degrees of belief, which are *subjective* mental states. Many Bayesians—*objective* Bayesians—employ the probability calculus primarily—sometimes exclusively—to represent or model other *objective* quantities, like confirmational, evidential, and logical support relations, objective chances, and relative frequencies. Thus, objective Bayesians hold that these quantities conform to the probability axioms in accord with a principle akin to (2) below. Many subjective Bayesians use probabilities to represent the aforementioned objective quantities too, but the use of probabilities to represent degrees of belief is distinctive of subjective Bayesianism.

decision theory, and statistics. Subjective Bayesianism's successes in illuminating and solving problems in these areas are well documented.[3] However, subjective Bayesianism is not a strongly unified movement. Subjective Bayesians often disagree about how exactly to characterize their view.[4] I will follow Easwaran (2011, p. 312) in characterizing subjective Bayesianism in terms of the following three commitments:

(1) there is kind of a gradational belief-like mental state, called "degree of belief";

(2) at any given time, one's degrees of belief ought to be faithfully representable with a probability function; and

(3) one's degrees of belief ought to change over time in light of new evidence according to the principle known as "conditionalization."[5]

Subjective Bayesians typically also accept some combination of several other probabilistic principles for degrees of belief, including the regularity principle, the principal principle, the principle of indifference, the reflection principle, and other alternatives to these principles.[6]

---

[3]See, e.g., Earman (1992, Chs. 3 & 4), Howson & Urbach (2006), Hájek & Hartman (2010).

[4]Some Bayesians hold that what is distinctive of Bayesianism is the centrality of Bayes theorem. Others think that it is the degree of belief interpretation of probability. See Weisberg (2011) for an account of the differences among Bayesians.

[5]According to conditionalization, assuming the representation of an agent's degrees of belief with a real-valued function, $cr(\cdot): \mathcal{L} \rightarrow \Re$, where $\mathcal{L}$ is a logical language, if the agent learns some claim, $A \in \mathcal{L}$, then the agent's degrees of belief, represented by $cr(\cdot)$ should *update* in light of this new information. If the agent's old degrees of belief were aptly represented by $cr_{old}(\cdot)$, then the agent's degrees of belief in light of learning $A$ should be aptly represented by $cr_{new}(\cdot) = cr_{old}(\cdot\,|A)$. Furthermore, if $cr_{old}(\cdot)$ is probabilistically coherent and $cr_{old}(A) \neq 0$, then $cr_{new}(\cdot)$ is coherent too, and $cr_{new}(A) = 1$. See Easwaran (2011, p. 314).

[6]See Lewis (1980) for his original statement of the principal principle. See van Fraassen (1984) for the original statement of the reflection principle. See Keynes (1921) for the principle of indifference, and see Jaynes (2003) for its modern reincarnation, the

Probabilism can be characterized as a commitment to the first two core tenets of subjective Bayesianism, that is, as the view that we have degrees of belief, and at any given time they ought to be faithfully representable as probabilities. Thus, probabilism constitutes the *synchronic* core of subjective Bayesianism, while conditionalization can be said to be its core diachronic principle.[7]

"Probabilism" is often used to refer both to a commitment to tenets (1) and (2), and to the latter principle. Hereafter, I will reserve "probabilism" to refer to the view constituted by the commitment to (1) and (2), and I will refer to (2) considered independently as "the probabilist norm."

To some readers familiar with probabilism and the notion of *degree of belief*, my statement of (1) and (2) above in terms of the representation of degrees of belief with numerical functions might look a little funny. For many probabilists take it as a foregone conclusion that degrees of belief are representable with numerical functions. Indeed, many probabilists seem to hold that degrees of belief somehow *are* numerical functions. But that's not how I think of degrees of belief. For my purposes in this dissertation, degree of belief is a kind of doxastic attitude in its own right, akin to the traditional epistemological notion of *categorical belief*. For my purposes at this point, I ask the reader to think of degree of belief as an attitude that the world is a particular way, an attitude which can be weaker or stronger rather than all-or-nothing. At this point, I make

principle of maximum entropy. For a thorough list of statements of the principle of regularity as well as critical discussion, see Hájek (MSb) and Hájek (2013).

[7]I should note that conditionalization is more controversial than probabilism. For a thorough discussion of alternate update rules, see Titelbaum (2013).

no assumptions about the relationships among degrees of belief and any other putative kinds of doxastic state: for example, I do not assume that categorical beliefs can be reduced to degrees of belief or *vice versa*. I will say more about these connections in later chapters, especially Chapters 2 and 4. Probabilism holds that our degrees of belief ought to be *faithfully representable* with probability functions.

Probabilism is an attractive view for several reasons. First, the notion of *degree of belief* often comes in handy for describing our doxastic states in particular cases. For example, our commitments to certain claims often seem to grow weaker or stronger in light of changes in evidence. Degrees of belief allow us to characterize these changes more precisely than the traditional notion of *categorical belief* alone. Second, probabilities seem to be apt for characterizing degrees of belief and their logic. The probability calculus can be viewed as a gradational extension of classical deductive logic, and it characterizes a kind of coherence akin to deductive consistency. Since many epistemologists accept deductive consistency as a kind of norm for categorical belief, the probabilist norm is an attractive analogue for degrees of belief.[8] Furthermore, as Titelbaum (2013, pp. 3-4) notes, probabilities are capable of capturing complex and varied evidential relationships. So, probabilities are very useful for representing how degrees of belief ought to accord with evidence. Finally, as a component of subjective Bayesianism, probabilism suggests illuminating positions on important problems in

---

[8]I do not mean here to endorse deductive consistency as a norm for categorical belief. Nor do I intend to suggest that the plausibility of the probabilist norm stands and falls with the plausibility of a consistency norm for categorical belief. See, e.g., Christensen (2004) and Easwaran & Fitelson (2015) for relevant discussion.

epistemology like the lottery and preface paradoxes, and the problem of (peer) disagreement.

Given probabilism's attractive features and track record, it is no surprise that probabilists are keen to apply their theory to every problem they can get their theoretical hands on. Indeed, many of the papers in the probabilist literature take a standard format: the author immediately characterizes probabilism with some variants of (1) and (2) followed shortly by an application of the theory to some particular problem by way of lots of formalism without much explicit interpretation of this formalism. This format makes a lot of sense: probabilism is a simple, powerful view, ripe for applications; it is fairly easy to grasp without a lot of explicit interpretation of the formalism.

While probabilism's simplicity, power, and intuitiveness are undoubtedly theoretical virtues, they have the side effect that the view is rarely articulated in sufficient detail.[9] Probabilists rarely take much time to explain what degrees of belief really *are*. They rarely present their probabilistic formalism in detail. And they don't spill much ink explaining *how* probabilism's formalism should be applied to particular cases, or how these applications should be interpreted to yield normative verdicts about degrees of belief. Fortunately, these words from Savage (1972, p. 1) still seem to hold true: "…here, as elsewhere, catastrophe is avoided, primarily because in practical situations common sense saves all but the most pedantic of us from flagrant error."

[9] Titelbaum (2013) is a notable exception. Indeed, much of this project is inspired by—and intended as response to—the early chapters of that book.

While the tenets of probabilism are not often articulated very fully or explicitly, the view is developed and revealed in its applications.

Despite its various attractive features and successful applications, probabilism is not without critics. It is subject to several important, well-known objections. Critics—often probabilists themselves—have objected to probabilism's precise numerical representation of degrees of belief. They have also objected to the excessive demandingness of the probabilist norm's requirement that our degrees of belief should be representable probabilistically. The probabilist norm seems to require us to be opinionated, with judgments about every logical combination of claims in our ken, and logically omniscient, certain of all logical truths. These requirements are far beyond our mnemonic and computational capacities.

Probabilists have developed several broad strategies for responding to these objections. For example, they've explained that the view's numerical precision is a kind of methodological idealization, a simplification for the sake of representational and computational simplicity and tractability. They've also explained that the view is a normative idealization as well: it is not meant to describe the way actual agents typically apportion their degrees of belief; rather, it is meant to characterize the notion of *ideal degree of belief* or some other such normative notion. These are just a couple of the probabilist strategies for avoiding objections.

In general, probabilists seem to be pretty well satisfied with these responses to the standard objections. When any of the standard objections threaten an application of the theory, probabilists acknowledge the appearance of the problem, deploy the appropriate response strategy, explain how the objection arises out of a naïve or unsophisticated interpretation or application of the view, and then they get on with their business.

The resulting dialectic is self-perpetuating. Probabilism's typical under-articulation leaves it open to the standard objections. Probabilists deploy their fleet of standard responses to these objections, but they do so piecemeal. The standard responses are individually plausible ways of deflecting or dissolving particular objections and filling in some of the details of the view. But it's not entirely clear that these responses can be employed coherently in unison to fill out probabilism's details in a plausible way. So when a probabilist deploys one of the standard responses to deflect one of the standard objections, it usually leaves room for the other standard objections to creep back in. It's kind of like trying to smooth out a bump in a rug. If you manage to smooth things out in one area, the problem crops up somewhere else. It's not clear that the probabilists can get the rug to lie flat.

I contend that probabilists must provide a unified articulation of probabilism's foundations in order to avoid the standard objections. To articulate their view's methodological foundations, probabilists must carefully articulate their formal system, its interpretation, and its domain of application.[10] Probabilists must specify whether their formalism is set-theoretic or logical, qualitative or quantitative, precise or

[10]The following aspects of probabilism's methodological foundations correspond to the elements of a well-articulated formal modeling framework characterized by Titelbaum (2013, p. 31).

imprecise, etc. While it is a pain to articulate these features of the view, these choices make a big difference in what probabilism amounts to. Once probabilists have specified their formalism, they must provide its standard interpretation. This means characterizing degrees of belief, which it is meant to represent, explaining how degrees of belief are represented in the formalism, and explaining how features of the representation can be applied to degrees of belief by explaining how the formalism characterizes some normative notion like *good*, *right*, or *rational degree of belief*, and by explaining how the formalism can be used to generate normative verdicts about degrees of belief in particular cases. Probabilists must also delimit their view's domain of application by specifying when it can and cannot be aptly applied. They must clarify whether their formalism is to be applied to represent agents' entire doxastic state or just local parts. And it would be helpful if they could characterize contexts in which probabilism may not be apt for representing degrees of belief.

An important part of specifying the interpretation of probabilism's formalism is specifying the normative foundations of the view. As I note above, probabilists must characterize the normative notion their formalism is supposed to capture. They must also carefully formulate the probabilist norm. They must specify its logical structure (whether it is wide or narrow scope, strict or slack, positive or negative polarity), and they must characterize its normative force (evaluative, deontic, or hypological). Furthermore, they should characterize the source of this normativity. This may mean more than simply appealing to one of the standard arguments for probabilism, like the

Dutch book argument or the epistemic utility arguments. These arguments should not only justify the probabilist norm, but also help to characterize it and its normative force.

My purpose in this dissertation is to show why probabilists need to better articulate the foundations of their view, to characterize what exactly it means to do this, and to provide an account of the core aspects of these foundations. As Savage (1972, p. 1) suggests, theory-building is unlike house-building insofar as foundations don't come first. However, I contend that probabilism has reached a level of theoretical maturity where its foundations can no longer be ignored. This dissertation is an attempt to demonstrate the need for an articulation of the foundations of probabilism and to defend one way to do so.

### 1.2 Chapter Outlines

This dissertation has four substantive chapters. Chapter 2 is devoted to presenting probabilism in detail. The purpose of this chapter is twofold: to present probabilism, the subject of this dissertation; and to demonstrate various ways in which probabilism still needs some filling in. I describe the most prominent proposals for fleshing out the view, and point out some unconsidered and undeveloped approaches. First I present common probabilist characterizations of *degree of belief*. Probabilists are committed to the existence of this doxastic state, and to its representability with real-valued probability functions. But they usually don't explain in enough detail what it means for agents to have degrees of belief, what the objects of degrees of belief are supposed to be, what it means for beliefs to come in degrees, how degrees of belief relate to other doxastic states, or why real-valued functions, probability functions in particular, are apt for representing degrees of belief. Next I consider the probabilist norm, according to which degrees of belief should obey—or conform to—the probability calculus. I present both set-theoretic and logical versions of the probability axioms, and I consider prominent views about what it means for degrees of belief to obey—or be representable as conforming to—these axioms. Finally, I present the three main classes of arguments that have been proposed to justify the probabilist norm: Dutch book arguments, representation theorem arguments, and epistemic utility arguments. I explain how each of these justifications also serves as partial characterization of probabilism, and I explain how each of them yields a slightly different characterization of degrees of belief.

Chapter 3 characterizes the main elements of the dialectic between probabilists and their critics. I focus on three main sources of criticism: probabilism's numerical representation of degrees of belief; its requirement that agents should be opinionated—that they should have degrees of belief about everything; and probabilism's logical omniscience requirement—that agents should be certain of all logical truths. I present objections to these aspects of probabilism, and describe probabilist responses to these objections. I show how this dialectic raises deep questions about the foundations of probabilism. I draw out of this dialectic several general strategies that probabilists use to respond to objections, and I show how these strategies amount to commitments to views about probabilism's normative and methodological foundations. I show how some of these strategies seem to constitute conflicting commitments about the nature of probabilism's foundations. Finally, I argue that probabilists need to articulate these foundations in a unified, systematic way in order to adequately address their critics' objections and demonstrate the co-tenability of their responses.

In Chapters 4 and 5 I sketch a unified account of probabilism's foundations. In Chapter 4 I focus on methodological foundations. First I briefly explain the virtues and vices of appealing to formal methods in general, and specifically those that come with appealing to probabilities to represent degrees of belief. I refer to this discussion throughout the remainder of the chapter in order to better explain the goals of applying probabilism's formalism and to show how to avoid certain pitfalls. Then I defend three main methodological positions. I defend a logical—rather than set-theoretic—formalism, on the basis that this approach is necessary for the appropriate modeling of non-ideal agents. I defend a fundamentally qualitative and comparative formalism, which allows us to better understand and avoid objections to probabilism's mathematical/numerical precision, and which allows us to state more intuitive and perspicuous connections between this formalism and norms for degrees of belief. I also defend a global—rather than merely local—approach to probabilist modeling, according to which probabilist modeling not only generates normative verdicts about particular cases, but also characterizes general connections between the probability calculus and norms for degrees of belief.

Chapter 5 is devoted to probabilism's normative foundations. I defend several normative positions. First, I lay out the apparatus of bridge principles—principles that connect the probabilist formalism with degree of belief norms—and I distinguish several different parameters that can be varied in the formulation of these principles. I defend a kind of normative pluralism, according to which there are a various different normative connections between the probability calculus and degrees of belief. But then I focus on elaborating a set of evaluative—rather than deontic or hypological—probabilistic coherence norms. Finally, I justify these bridge principles on the basis that while none of the standard arguments for the probabilist norm suffice on their own, together these arguments, along with subjective Bayesianism's many successes, provide strong support for normative connections between the probability calculus and degrees of belief.

## 1.3 Terminological & Notational Conventions

To allay future confusion, I want to announce (and in some cases reiterate) some terminological and notional conventions I will employ in the remainder of the dissertation. It's a little awkward to present these conventions out of context, and I will remind the reader of these conventions at various points in later chapters. However, the literature on probabilism is so full of conflicting uses of technical terms and formalism that I think that I should sort out some of these issues up front.

### 1.3.1 Probabilism & the Probabilist Norm

The subject of this dissertation is probabilism. By "probabilism" I mean the view that posits the existence of degrees of belief and invokes probability theory as a kind of logic and a source of norms for these degrees of belief. Some probabilists also use "probabilism" to refer to the norm or principle that degrees of belief ought to conform to the axioms and theorems of the probability calculus. To avoid confusing the view with the norm, I will refer to the norm as "the probabilist norm."

### 1.3.2 Doxastic Attitudes

In this dissertation I talk a lot about doxastic attitudes, mental states about the way the world is (as opposed to how we desire, intend, or imagine the world to be). In particular, I talk a lot about three main kinds of doxastic attitudes: degree of belief, categorical belief, and comparative confidence. These doxastic attitudes are all known by many different names in the literature, and there are many different views about how these attitudes relate to each other. Below I briefly explain the terms I will use to refer to these attitudes.

I focus primarily on degree of belief. For, as I just noted, one of the core tenets of probabilism is the claim that there is a gradational belief-like mental state called "degree of belief." Probabilists use various terms to refer to degrees of belief in addition to the term "degree of belief." They use such terms as "credence," "degree of credence," "degree of confidence," "level of confidence," and "partial belief." While some probabilists might acknowledge subtle distinctions among the notions these terms pick out, I do not. I take all of these terms to pick out a kind of gradational doxastic state akin to the traditional notion of *categorical belief*. To avoid confusion, I will try to avoid variation and stick to using "degree of belief" to refer to this mental state. However, in some cases, it is more natural to talk about degrees of belief in terms of confidence or degrees of confidence. So, I diverge from "degree of belief"-talk from time to time. I may also slip up from time to time, and slip into using "credence"-talk, etc. Sorry in advance.

In addition to degree of belief, I will also discuss categorical belief and comparative confidence. Categorical belief is a categorical, all-or-nothing attitude about the way the world is. Comparative confidence is a comparative judgment of the relative strength of an agent's confidence in two claims. Like degree of belief, these attitudes are also often referred to by several names in the literature. Categorical belief is often called "belief," "full belief," "binary belief," "flat-out belief," "all-or-nothing belief," and more. While comparative confidence is not discussed as often as categorical belief or degree of belief, it is also known by multiple names. It is sometimes referred to as "comparative probability" or "qualitative probability."

Though I have introduced different terms for degree of belief, categorical belief, and comparative confidence, I do not assume that these are genuinely distinct kinds of doxastic attitudes. As I note above, there are many different views about how these attitudes relate to each other. For example, many probabilists hold that categorical beliefs are simply a species of degree of belief. Others hold that both categorical belief and degree of belief can be reduced to comparative confidence judgments. In the end, I will hold that degree of belief and comparative confidence are really the same kind of gradational doxastic attitude.[11] However, at this point, I want to be able to talk about these attitudes as though they are distinct, whether or not they actually are.

### 1.3.3 Formal Representations of Doxastic Attitudes

In addition to talking a lot about the aforementioned doxastic attitudes themselves, I will also talk a lot about their formal representations. In particular, I'll talk a lot about the formal representation of degrees of belief with real-valued functions. So, while I will

[11] I will not take a stand on the relationship between categorical belief and degree of belief.

treat the terms "degree of belief" and "credence" as co-referential in what follows, I will use "credence function" rather than "degree of belief function" to refer to the mathematical functions probabilists use to represent degrees of belief. This is because I want to be very clear that degrees of belief are distinct from the functions that represent them. Degrees of belief are mental states that are not necessarily numerical; credence functions are functions from expressions of a formal language (sentences or sets) to real numbers between 0 and 1. I hope that adopting this terminological strategy will help both me and my reader to avoid conflating degrees of belief and credence functions.

Likewise, I want to avoid conflating the other doxastic states, categorical belief and comparative confidence, with their formal representations. I will not discuss the formal representations of categorical belief in this dissertation, so I will set aside that issue. However, I will discuss the formal representation of comparative confidence judgments with what I call "comparative confidence relations." The comparative confidence relations ($\succcurlyeq$, $\succ$, $\approx$, and $\sim$) are relations in a formal language. They are designed to represent agents' confidence comparisons, but they may not necessarily do so perfectly. There may be aspects of confidence comparisons that are not faithfully represented by comparative confidence relations, and there may be features of the comparative confidence relations that comparative confidence judgments lack. Thus, I want to avoid conflating comparative confidence judgments with the comparative confidence relations that represent them.

### 1.3.4 Logical & Mathematical Formalism

I assume a familiarity with some logic and mathematics. I will not attempt to survey those assumptions here. However, I do want to present the logical and mathematical notation I will employ, which is summarized in the following table.

Table 1: Logical & Mathematical Symbols

| Symbol | Interpretation |
|---|---|
| $A, B, C, \ldots, A_1, B_1, \ldots$ | Metalinguistic variables for sentences of a formal language |
| $\neg$ | Negation |
| $\wedge$ | Conjunction |
| $\vee$ | Disjunction |
| $\rightarrow$ | Material conditional |
| $\leftrightarrow$ | Material biconditional |
| $\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots$ | Sets |
| $\mathrm{A} - \mathrm{B}$ | Set-theoretic difference (the set of members of $\mathrm{A}$ not in $\mathrm{B}$) |
| $\overline{\mathrm{A}}$ | Set-theoretic complement of $\mathrm{A}$ (in the context of a set of all elements of interest, $\Omega$, $\overline{\mathrm{A}} = \Omega – \mathrm{A}$) |
| $\cap$ | Intersection |
| $\cup$ | Union |
| $\in$ | Set membership |
| $\subset$ | Subset relation |
| $\subseteq$ | Proper subset relation |
| $cr(\cdot)$ | Unconditional credence function |
| $cr(\cdot \mid \cdot)$ | Conditional credence function |
| $pr(\cdot)$ | Unconditional probability function |
| $pr(\cdot \mid \cdot)$ | Conditional probability function |
| $\succ$ | Strict comparative probability relation |
| $\succcurlyeq$ | Weak comparative probability relation |
| $\approx$ | Comparative probability indifference relation |
| $\sim$ | Comparative confidence indeterminacy relation |

## 2 Probabilism

Here's a pretty typical two-tenet characterization of probabilism:

1. beliefs come in degrees, and
2. degrees of belief ought to conform to the probability calculus.[12]

According to this characterization, probabilism is a commitment to the existence of a particular kind of doxastic state, degree of belief, and to a norm for degrees of belief.

Obviously, this characterization is incomplete. It leaves a lot of room for interpretation and further characterization. The first tenet tells us that we have degrees of belief. But it doesn't tell us what degrees of belief *are*, what their objects are, what it means for them to come in degrees, how they relate to other doxastic states (if there are any), how they should be represented formally, or why we should think there are degrees of belief in the first place. And the probabilist norm tells us that the axioms of the probability calculus are normative constraints on the apportionment of degrees of belief. But it doesn't tell us *how* the probability axioms, which specify a kind of real-valued function, serve as norms for a kind of doxastic state. Probabilists must explain how degrees of belief can be represented with real-valued functions *and* how the axioms and theorems of the probability calculus can be applied to degrees of belief as norms.[13] For while the probabilist norm sounds innocuous to the initiated, it can sound like a category mistake to those who are not steeped in Bayesian lore.

---

[12]For similar statements of probabilism see, for example, Eriksson & Hájek (2007, p. 183), Hájek (2008, pp. 793-794), Kaplan (2010, p. 42), Joyce (1998, pp. 575-576; 2009, pp. 263-264), Zynda (2000, p. 45), and Pettigrew (2011).

[13]This second task is often ignored by probabilists. Titelbaum (2013) stresses this point.

Thus, this common characterization of probabilism leaves many questions unanswered. That's not too surprising, since it's stated in just one sentence. But the problem, as I note in the introduction, is that probabilists rarely answer these questions very explicitly. They tend to focus instead on applications of the view. And while probabilists *do* give various arguments designed to justify the probabilist norm, and these arguments *do* begin to fill in the details of the view, they do not do so very explicitly or completely. Since probabilism is the subject of this dissertation, my purpose in this chapter is to present the probabilist thesis in detail. But, since probabilism is rarely presented in much detail, my approach will be to consider each of probabilism's two core tenets, pointing out the ways in which the view is under-articulated, and presenting some of the usual options for filling it out and justifying it.

§2.1 is devoted to characterizing degrees of belief. §2.2 is a survey of the main reasons that probabilists posit degrees of belief. §2.3 is an explanation of the probabilist norm, including a presentation of the probability calculus and its typical presentation as a source of norms for degrees of belief. §2.4 is a survey of the main arguments probabilists give to justify the probabilist norm, including Dutch book arguments, representation theorem arguments, epistemic utility arguments, and what I call "normative triangulation" arguments.

**2.1 What Are Degrees of Belief?**

Probabilism's first tenet posits the existence of a gradational doxastic state called "degree of belief." But, as I note above, it doesn't tell us what it means to have this particular kind of doxastic attitude, what its objects are, what it means for it to come in degrees, how it relates to other kinds of doxastic states that are commonly posited in epistemology, or how exactly it should be represented formally. In this section, I'll pursue these questions in order, announcing the different options that probabilists have for answering them.

Before I continue, I want to expand on some comments I made in the introduction intended to allay confusion about my conception of degree of belief. As I note in §1.3, degree of belief is a concept that goes by many names: "degree of belief," "credence," "degree of credence," "partial belief," "degree of confidence," "level of confidence," and more. Different authors employ different terms for various reasons, and some acknowledge subtle distinctions among these terms and among the concepts to which they refer.[14] In this dissertation, I will treat these terms as co-referential.

To some probabilists, it may sound odd to talk about degrees of belief as mental states that can be considered independently of the probability calculus, and to raise the question, as I do above, of how degrees of belief relate to probabilities. This is because the notion of degree of belief emerged out of the subjective interpretation of probability (Titelbaum 2013, p. 16). On this conception, it doesn't make much sense to ask whether degrees of belief should be represented formally as probabilities, or whether the probability axioms are appropriate norms for degrees of belief.

As I explained in the introduction, this is not how I think of degree of belief. I think that we can consider the notion of degree of belief independently, without reference to

[14]See, for example, Sturgeon (2010), who prefers to trade in terms of confidence, and who reserves "credence" to refer to sharp numerical subjective probabilities.

probability.[15] I think of degree of belief as a legitimate doxastic attitude in its own right, akin to the traditional epistemological notion of belief. I think that real human agents have degrees of belief, just as we are typically thought to have beliefs. I do not reserve "degree of belief" to refer to the doxastic attitudes of ideally rational agents. I do not think of degrees of belief as necessarily explicitly numerical, precise, or complete. I see my view of *degree of belief* as continuous with traditional epistemological theorizing about *belief*. I think it is meaningful to ask why and how degrees of belief should be represented formally. And I think it is meaningful to ask whether probability functions are apt for this purpose, and whether probabilistic coherence is an appropriate norm for degrees of belief. So, I ask the reader to set aside the conception of degree of belief as the subjective interpretation of probability, and, in what follows, I invite the reader to consider afresh what exactly degrees of belief are, whether and how they should be represented formally, and which norms are appropriate for them.

### 2.1.1 What Does It Mean to Have Degrees of Belief?

Part of understanding what degrees of belief *are* is understanding what it means to *have* them. Degrees of belief are a kind of doxastic state or attitude.[16] The notion *doxastic attitude* seems to be a precisification of the folk psychological notion of *belief*, which is, roughly, a mental attitude about the way the world is. Doxastic attitudes are distinguished from so-called "conative," desire-like attitudes by their "mind-to-world"

[15]Here I echo Titelbaum (2013).

[16]I do not distinguish between doxastic states and doxastic attitudes. Some use "doxastic state" to refer to all of one's doxastic attitudes at a time. I call that one's "total doxastic state."

rather than "world-to-mind" direction of fit.[17] That is, doxastic attitudes are oriented at fitting the way the world is, whereas conative attitudes are oriented at changing the world to fit themselves. Epistemologists and philosophers of mind have developed several theories of what it means to have doxastic attitudes—to have beliefs *qua* mental states. These theories include eliminativism, instrumentalism, functionalism, representationalism, dispositionalism, interpretationalism, and primitivism. Probabilists can adapt these theories to *degree of belief*, giving them several options for accounts of what it means to have degrees of belief. In this subsection, I will survey these options and note connections with common probabilist characterizations of *degree of belief*.[18]

Eliminativism about doxastic states is the view that there are no doxastic states, and that, accordingly, the notion *doxastic state* should be abandoned. Thus, eliminativists must deny the existence of degrees of belief. Since probabilism's first tenet is a commitment to the existence of degrees of belief, and since probabilism is predicated on the theoretical utility of *degree of belief*, eliminativism is incompatible with probabilism.

---

[17] According to Humberstone (1992), the distinction between mind-to-world and world-to-mind direction of fit is due to Anscombe (1957). The direction of fit terminology was popularized by Smith (1987, 1988, 1994). While this terminology is evocative and, thus, helpful in distinguishing doxastic and conative attitudes, its use has been criticized (see Sobel & Copp 2001). I employ this terminology here and later in the dissertation to provide a rough characterization of doxastic attitudes. I do not fully endorse the "direction of fit" theory of attitudes.

[18]The characterizations of the accounts of the nature of degree of belief surveyed below draw heavily on Eriksson & Hájek (2007) and Schwitzgebel (2011).

While probabilism is incompatible with eliminativism, it is compatible with at least some forms of instrumentalism. Soft instrumentalism about doxastic states is the view that there are doxastic states, but that they are not robustly real. As Schwitzgebel (2011, §1.5) notes, a soft instrumentalist about doxastic states might compare them to entities like the equator, centers of gravity, or the average American. There is no thin red line painted around the earth, but it is nonetheless meaningful to explain the location of Quito, Ecuador by appeal to the equator. The equator is less real than Quito, but more than a mere fiction. Likewise there may not be any concrete phenomena that can be called "degrees of belief," but degrees of belief might nonetheless exist and *degree of belief* might be theoretically useful. Hard instrumentalism about doxastic states is the view that—strictly speaking—there are no doxastic states, but that it can be theoretically useful to talk as if there are. For the same reason that probabilism is incompatible with eliminativism, it seems to be incompatible with hard instrumentalism. I think, however, that probabilism's first tenet could be modified to make it compatible with hard instrumentalism: rather than positing the existence of degrees of belief, it might merely state that it is useful to talk *as if* there are degrees of belief. Depending on the details, each of the views presented below can be given an instrumentalist interpretation. Dispositionalism and interpretationalism are especially amenable to instrumentalist interpretations.[19]

[19]Indeed, you might even think that these views require at least a soft form of instrumentalism.

Functionalism about doxastic attitudes is the view that what makes something a doxastic attitude of a certain kind are its actual, potential, or typical causal relationships, especially causal relationships with other mental states, sensory stimulation, and observable behavior (Schwitzgebel, 2011, §1.4). Functionalism can be further characterized in many different ways depending on which causal roles we treat as distinctive of the kind of doxastic state in question. Some examples of these characteristic causal roles are relations to the tendency to assert, the tendency to be formed in response to perceptual experiences, the tendency to infer the logical consequences of the content of one's doxastic states, and the tendency to produce actions when combined with certain desires and intentions (Schwitzgebel, 2011, §1.4). The other accounts of the nature of doxastic states surveyed below can all be viewed as versions of functionalism.

Eriksson & Hájek (2007, p. 196) characterize Ramsey (1964) as the first functionalist probabilist. Ramsey held that degrees of belief are the things, which when combined with desires—or more precisely utilities—serve as the bases of our actions. Ramsey's view also has affinities with interpretationalism, as I will discuss below.

Representationalism is the view that having a doxastic attitude toward a doxastic object is just a matter of having a particular mental representation that captures the content of the doxastic object (Schwitzgebel, 2011, §1.1).[20] These representations play

[20]Representationalism in this sense is not to be confused with characterizations of degrees of belief in terms of representation theorem arguments for probabilism, which have close ties to interpretationalism.

a causal role in mental operations (like inferences and decisions) and the production of behavior (like language use and all manner of other actions). Representationalists offer different proposals about exactly what these mental representations are and how they must be stored in the mind in order to count as doxastic—rather than, for example, conative or imaginative—representations. Fodor (1975), for example, suggests that they are sentences in a species-wide internal "language of thought." These representations count as doxastic attitudes when they are stored in one's "belief box."

It's not entirely clear how the representationalist probabilist would account for the *gradational* structure of degrees of belief. They might hold that mental representations bear markers or valences of some kind that indicate strength of confidence. Or they might hold that the "belief box" is structured, with stronger degrees of belief residing in one region, and weaker degrees of belief residing in another. They might even think of degrees of belief as mental representations that are explicitly probabilistic. For example, on this view, a strong degree of belief that Elvis was a member of the Beatles might be constituted by a sentence in the language of thought equivalent to "it is highly probable that Elvis was one of the Beatles."[21] The representationalist probabilist would have to

[21]For arguments against placing the gradational structure of degrees of belief into the content of doxastic objects, see Christensen (2004, pp. 18-20) and Eriksson & Hájek (2007, pp. 206-207). These arguments do not concern representationalist theories of doxastic states in particular. Rather, they criticize the general strategy of treating degrees of belief as categorical beliefs with probabilistic content.

modify the standard representationalist accounts of belief in order to fully specify their[22] view.

Dispositionalism is the view that having a particular doxastic attitude is a matter of having particular mental or behavioral dispositions (Schwitzgebel, 2011, §1.2). Dispositionalism comes in a few different forms. It is often spelled out as the view that being in a particular doxastic state is just a matter of the pattern of one's actual and potential behavior. On this kind of view, to have a particular degree of belief might be cashed out as the tendency to say certain things, exhibit surprise in particular circumstances, perform certain actions, and so on. Traditional behaviorism is a degenerate form of dispositionalism, according to which having a particular doxastic attitude is just a matter of one's actual observable behavior. But not all forms of dispositionalism are purely behavioristic. More liberal forms of dispositionalism also explain doxastic states in terms of dispositions toward having other mental states—possibly even other doxastic states. On this sort of view, one might count as having a particular degree of belief if one is disposed to speak certain ways, act certain ways, or even to hold certain other degrees of belief. Dispositionalism seems to be well-equipped to account for the gradational structure of degrees of belief, for dispositions are also matters of degree. An agent with a strong degree of belief that $A$ will have a strong disposition to utter "$A$" and to have strong degrees of belief in $A$'s logical consequences,

[22] I use "their" as the generic possessive pronoun here rather than "his," "her," "his/her," or "his or her" for the sake of gender inclusiveness on the basis that many people do not identify with either of those genders. I follow this principle throughout the dissertation.

a weak disposition to express surprise at $A$, and so on. One with a weak degree of belief that $A$ will have a weak disposition to utter "$A$," a strong disposition to exhibit surprise at $A$, and so on.[23]

The well-known betting interpretation of degrees of belief has strong affinities with dispositionalism. De Finetti (1974) famously advocated a behaviorist operational definition of "degree of belief" in terms of actual betting behavior. On his view, one's degree of belief in a claim, $A$, is the ratio $\frac{\$x}{\$y}$ that one deems fair for buying or selling a bet of $\$x$ on $A$ for a $\$y$ stake. De Finetti's operationalism quickly went out of style, but the characterization of degrees of belief in terms of betting has been popular among probabilists in various iterations for a long time. The idea is that all of our decisions are gambles on outcomes, so rational action requires rational bets based on our belief-strengths and our desire-strengths (our utilities). Our best current accounts of decision theory follow this line of thought.

Like some forms of dispositionalism, interpretationalism also accounts for the nature of doxastic states in terms of behavior. However, interpretationalism also includes an additional appeal to a theoretical interpretive framework (Schwitzgebel, 2011, §1.3). Whereas a behaviorist would hold that doxastic attitudes *just are* a subset of one's actual behaviors or behavioral dispositions, the interpretationalist holds that one's doxastic attitudes are theoretical entities posited to explain one's behavior. On this view, what it means to be in a doxastic state is to exhibit particular behaviors (or

[23]The dispositionalist probabilist will, of course, have to account for how having low degree of belief that $A$ differs from having no such degree of belief.

behavioral dispositions) that can be predicted and explained by attributing the doxastic attitude in question, along with other doxastic, conative, and—perhaps—affective states. So, an interpretationalist probabilist would say that degrees of belief are theoretical entities posited to explain behavior. For example, one's strong degree of belief that it will rain today is a doxastic state posited to explain one's wearing a rain coat and carrying an umbrella.

Interpretationalism has also long been popular among probabilists. Traces of the view can be found in Ramsey (1964), and it can be found more explicitly in Lewis (1974) and Maher (1993). Like dispositionalism, interpretationalism is often associated with representation theorem arguments for probabilism, which I'll discuss in §2.4.2.

Depending on which of these views probabilists hold, they might be characterized as instrumentalists or realists about degrees of belief.[24] Since probabilism's first tenet is a commitment to the existence of degrees of belief, probabilists' choice among these accounts partially determines what kind of existential commitment this is. Representationalism seems to require a realism about degrees of belief. But dispositionalism and interpretationalism seem to be amenable to soft instrumentalist interpretations, according to which degrees of belief are real. The nature of this existential commitment will have important consequences for the plausibility of various aspects of probabilism, including its formal representation scheme, and the probabilist norm itself.

[24]Since probabilism's first tenet asserts the existence of degrees of belief, it wouldn't make sense for probabilists to be eliminativists about degree of belief.

Viewed as analyses or explications of *degree of belief*, functionalism, representationalism, dispositionalism, and interpretationalism all face serious objections.[25] Primitivism is an alternative approach to characterizing degrees of belief. Primitivism about degrees of belief is the view that *degree of belief* is a primitive concept—a concept that cannot be faithfully reduced to, or analyzed or explicated in terms of other concepts.[26] While primitivism about degrees of belief denies that *degree of belief* can be analyzed or explicated, it does not deny that it can be *characterized*. On this view, we can say many interesting and fruitful things about degrees of belief, without analyzing *degree of belief*. Indeed, the arguments for probabilism can be viewed as providing such characterizations. As Eriksson & Hájek (2007, p. 209) note, some of the arguments for probabilism (to be discussed in §2.4) seem to assume primitivism about degrees of belief. The difficulties of analyzing degree of belief, in turn, serve as evidence that it is a natural primitive. So do the fruits of subjective Bayesianism.

*Prima facie*, there's nothing in the tenets of probabilism that explicitly favors any of these views. However, probabilism's plausibility depends on which view of the nature of degrees of belief probabilists adopt. Each view has theoretical merits and demerits of its own. But, additionally, each view has implications for the plausibility of other aspects

[25]See Eriksson & Hájek (2007) for a thorough account of these problems.

[26]I distinguish between conceptual analysis and explication as follows. The analysis of a concept is the articulation of necessary and sufficient conditions for satisfying that concept. Some concepts, however are vague or ambiguous, making them impossible to analyze in this sense. Such cases call for explication, the articulation of a partially stipulated set of necessary and sufficient conditions that specify a precise concept useful for some theoretical purposes, though not apt for capturing the full meaning of an everyday concept.

of probabilism, such as its formal representation scheme and its justification for the probabilist norm. I will say more about these connections in §2.4, when I present the various arguments for the probabilist norm.

### 2.1.2 What Are the Objects of Degrees of Belief?

Just as probabilists can choose from several theories of what it means to have doxastic attitudes, they can also choose from among several theories of the objects of these attitudes. The list of candidates is long: propositions (centered or uncentered), propositions under descriptions, epistemic possibilities, doxastic possibilities, natural language sentences, sentences in a language of thought, and more. Probabilism's core tenets don't clearly nominate or exclude any of these views. Each approach has its own merits and demerits. I will not attempt to survey all of these options here. I will focus on the views most popular among probabilists: propositions, and epistemic and doxastic possibilities. I will also present an approach, due to Titelbaum (2013), which allows us to model the objects of degrees of belief without taking a firm stand about what exactly they are. The underlying conception of doxastic objects is important for assessing probabilism's formal representation scheme and some of its apparent normative requirements, such as the logical omniscience requirement.

Propositions are the most commonly proposed doxastic objects. Traditional epistemologists often claim that belief is categorical assent to the truth of a proposition. Likewise, probabilists often claim that degree of belief is the strength of one's confidence in the truth of a proposition.[27] Most philosophers agree on a few basic features of propositions: they are the primary bearers of truth value, the contents expressed by declarative sentences (a.k.a. “claims” or “statements”) in context, the referents of that-clauses, and the objects of attitudes shared among agents (McGrath 2012). But beyond this core characterization, there are several competing accounts of exactly what propositions are. The two most prominent views of propositions are the possible worlds approach and the structuralist approach.

According to the possible worlds approach, a proposition is the set of possible worlds in which it is true. For example, the proposition *Patrick Epley lives in Norman, OK, on July 13, 2014* is the set of possible worlds in which I, Patrick Epley, live in Norman, OK, on July 13, 2014. It is widely agreed that possible worlds are maximally consistent ways the world could be. But, as usual, there are disagreements over the details.[28]

Structuralism holds that propositions are structures composed out of individuals, properties, and relations.[29] On this approach, *Patrick Epley lives in Norman, OK on July*

[27]Some even think of degrees of belief as numerical truth-value estimates, where true is 1 and false is 0. See Joyce (1998, p. 587) who attributes this view to Jeffrey (1986). Joyce (2005, p. 155) takes the somewhat weaker view that degrees of belief help us estimate truth values of propositions.

[28]Lewis (1986) holds that they are concrete entities as real as the actual world. Others hold that they are abstract objects of various kinds (for example, Adams (1974), Fine (1977) and Plantinga (1974, 1976)). Still others think that they are combinations of so-called metaphysical simples (for example, Armstrong (1989)). See Menzel (2013) for an overview.

[29]See King (2012) for an overview.

*13, 2014* is understood as a structure composed of me, Norman, OK, July 13, 2014, and the inhabitance relation.

Despite their popularity, propositional accounts of doxastic objects have a few well-documented problems. Most of these problems stem from propositions' coarseness of grain; propositional accounts of doxastic objects don't allow us to make fine enough distinctions between doxastic objects. For example, they don't make proper sense of doxastic attitudes towards logical truths and other necessary truths, attitudes about objects that are known under multiple descriptions, and self-locating attitudes.

Construed as sets of possible worlds, propositions are ill-suited to being the objects of doxastic states about logical truths and other necessary truths because all necessary truths express the same proposition, namely the set of all possible worlds. Thus, on this view, doxastic attitudes toward necessary truths are indistinguishable. But this conflicts with the common intuition that, for example, a belief that $2 + 2 = 4$ is importantly different from a belief that $29 \times 232 = 6{,}728$. Construed structurally, however, propositions appear to be better suited to being doxastic objects. Whatever the constituents of the propositions expressed by "$2 + 2 = 4$" and "$29 \times 232 = 6{,}728$" are, they are plausibly different. But this raises the question of what exactly these constituents are.

Neither the possible worlds approach nor the structural approach makes propositions especially well-suited to being the objects of doxastic attitudes about individuals known under multiple descriptions. That is, they both seem to fall afoul of Frege's puzzle (Frege 1997). For, on both accounts, attitudes toward *Hesperus is*

*Hesperus* and *Hesperus is Phosphorus* are attitudes toward the same proposition. On the possible worlds account, this proposition is the set of possible worlds in which the planet Venus is self-identical. On the structural account, it is a structure that predicates self-identity of that planet. In both cases, an attitude toward *Hesperus is Hesperus* is an attitude toward *Hesperus is Phosphorus*.

And neither account of propositions accords well with propositions being the objects of self-locating attitudes. For I might have a strong degree of belief that I live in Norman, OK, but a very weak degree of belief that Patrick Epley lives in Norman, OK, because I might have a weak degree of belief that I am Patrick Epley.[30] The possible worlds account of propositions can't make sense of my degree of belief. For "I live in Norman, OK" (written or spoken by me) and "Patrick Epley lives in Norman, OK" express the same proposition, namely the set of possible worlds in which I, Patrick Epley, live in Norman, OK. Likewise, the structural account of propositions also fails, for in both cases, the object of my divergent degrees of belief is a structure composed of me, Norman, and the inhabitance relation.

These problems motivate some probabilists to posit centered propositions as the objects of doxastic attitudes.[31] The centered propositions approach is an outgrowth of the possible worlds conception of propositions. Whereas a traditional, uncentered

---

[30]This example is adapted from Huber (2014).

[31]Lewis (1979) is the *locus classicus* for discussion of centered propositions. Lewis attributes the idea to Quine (1969). See Arntzenius (2003), Bostrom (2007), Bradley (2012), Elga (2000), Meacham (2008), and Titelbaum (2013) for discussion of centered propositions as the objects of degree of belief.

proposition is understood as a set of possible worlds, a centered propositions is a set of possible worlds indexed to an individual at a time. Some think of centered propositions as properties of individuals. One's belief in a centered proposition is a belief that the actual world centered on oneself is in the set of centered worlds that constitutes the centered proposition.

In the example above, we can make sense of my divergent degrees of belief that I live in Norman and that Patrick Epley does by appeal to centered propositions. The object of my strong degree of belief that I live in Norman is the centered proposition constituted by the set of possible worlds centered on me now in which I live in Norman. The object of my weak degree of belief that Patrick Epley lives in Norman is the centered proposition constituted by the set of all centered worlds—regardless on which individual and time they are centered on—in which Patrick Epley lives in Norman. Because centered worlds are indexed to time-slices of individuals, centered propositions serve as finer-grained doxastic objects than their uncentered cousins.

But centered propositions have problems of their own when construed as doxastic objects. *Prima facie*, they don't help us make sense of attitudes toward logical truths, and they don't help us solve Frege's puzzle. Furthermore, centered propositions are too fine-grained for some purposes. Taking them as the objects of degrees of belief conflicts with other elements of the Bayesian approach. While doing so doesn't much affect probabilism (Lewis, 1979, p. 534), it does affect Bayesian degree of belief updating (Huber 2014, §1.2). For the centered propositions that serve as the objects of one's self-locating doxastic attitudes will shift as one travels through spacetime. Since most probabilists endorse other norms within the Bayesian package, including Bayesian degree of belief updating, this problem makes centered propositions less plausible as doxastic objects.

Some philosophers have suggested that doxastic objects are sets of epistemic or doxastic possibilities—rather than logical or metaphysical possibilities.[32] Intuitively, the epistemic possibilities for an agent at a time are the possibilities that are not ruled out by the agent's store of knowledge at that time, and an agent's doxastic possibilities at a time are the possibilities that are not ruled out by their doxastic states at that time. These two approaches are initially attractive because they seem to avoid Frege's puzzle and issues about doxastic attitudes toward logically and metaphysically equivalent claims. For, if an agent doesn't know (believe) that Hesperus is Phosphorus, then *Hesperus is Hesperus* and *Hesperus is Phosphorus* may be different epistemic (doxastic) possibilities for the agent. So, the agent can have different doxastic attitudes towards *Hesperus is Hesperus* and *Hesperus is Phosphorus*. Similarly, for all the agent knows (believes) it may be impossible that $2 + 2 \neq 4$, but nonetheless possible that $29 \times 232 \neq 6{,}728$.

There are two main flaws with the epistemic and doxastic possibility accounts of doxastic objects. First, the notions of *epistemic* and *doxastic possibility* are not yet made precise, leaving several important questions unanswered. For example, what exactly does it mean for a possibility to be ruled out by your store of knowledge or your doxastic

[32] See Hacking (1967), Bjerring (2010), Chalmers (2011, 2011a), and Easwaran (2011, 2011a) for discussion of these approaches.

attitudes? And what kinds of possibilities are these that are ruled in or out by our knowledge or doxastic states?

Second, epistemic and doxastic possibility approaches assume a form of logical omniscience. This flaw is common to all set-theoretic accounts of doxastic objects. The problem is that sets are purely extensional. So, two sets of possibilities are identical so long as they are composed of the same members. Thus, complex possibilities—set-theoretic combinations of possibilities—are identical so long as they are composed of the same basic elements. Thus, for example, the sets $\mathrm{A} \cap \mathrm{B}$ and $\overline{\overline{\mathrm{A}} \cup \overline{\mathrm{B}}}$ are identical, though they are formed by different set-theoretic operations. In short, these approaches cannot account for agents who are set-theoretically non-omniscient. I'll say more about this issue in Chapter 4.

In response to difficulties with each of the most popular accounts of doxastic objects, Titelbaum (2013, pp. 35-37) sets aside the issue of exactly what the objects of our doxastic attitudes really are. Instead, he proposes to model degrees of belief *as if* they are about the truth of "claims" (declarative natural language sentences) in context. As Titelbaum (2013, p. 35) puts it, modeling doxastic objects this way provides an "access point" to all plausible theories of what doxastic objects really are. For whatever they really are, doxastic objects underwrite our ability to invest confidence in the truth of claims in context. This is why we can elicit someone's degree of belief in $A$ by asking "how confident are you that '$A$' is true in the present context?" (Titelbaum 2013, p. 37).

Modeling degrees of belief as if they are about the truth of claims in context allows us to avoid the problems with the popular accounts of doxastic objects. For claims in context are appropriately fine-grained because they are tied to the syntax of natural language sentences. Thus, this approach allows us to distinguish the doxastic objects modeled by "Hesperus is Hesperus" and "Hesperus is Phosphorus," "$2 + 2 = 4$" and "$29 \times 232 = 6{,}728$," "I live in Norman" and "Patrick Epley lives in Norman," etc. Indeed, claims in context may even be too fine-grained. For example, "Danica gave George the wrench" and "the wrench was given to George by Danica" model different doxastic objects on this approach. But fine grain is better than coarse grain in this context.[33]

### 2.1.3 *Degrees* of Belief?

What does it mean for a doxastic attitude to be gradational—to come in degrees? This question begets several others. To start with, what is the alternative? The obvious comparison is to the categorical doxastic attitudes that have long been the primary focus of epistemologists, namely belief, disbelief, and suspension of judgment. These doxastic attitudes are binary: you believe a claim or not; you disbelieve it or not; you suspend judgment on it or not.[34] They do not come in degrees. So, degrees of belief are not

[33] See Titelbaum (2013, pp. 35-37) for further discussion of the features and flaws of this approach.

[34] These categorical doxastic states also seem to be exclusive. For example, it seems that you can't both believe and disbelieve that Biden is French. However, these attitudes are not exhaustive in that there may be claims towards which you have no attitude whatsoever. We should not confuse suspending judgment on a claim with having no attitude toward the claim. Until you read this sentence, you probably had no attitude about whether I ate Chilean seabass for lunch on May 1st, 1993. But now that I've broached the subject, hopefully you've suspended judgment.

categorical (binary) like these other doxastic attitudes. They come in different grades of strength.[35]

How many grades of confidence are there? There might be just a few, akin to the Likert scale: strong belief, belief, equivocal confidence, disbelief, and strong disbelief.[36] Or there might be a larger finite number of levels of confidence. Or there might be infinitely many—countably or uncountably many—grades of confidence. As I've mentioned above, and as I'll explain in more detail in §2.1.5, probabilists typically represent degrees of belief with real numbers between 0 and 1. There are uncountably many such numbers. So, if the probabilist representation of degrees of belief with real-valued functions is not overly precise, this suggests that there are, in principle, uncountably many grades of confidence. However, as we'll see in Chapter 3, many probabilists concede that their formal representation of degrees of belief *is* overly precise in various ways. Regardless of whether there are finitely or infinitely many grades of confidence, there seem to be *many* such grades that allow us—in at least some cases—to make fine-grained distinctions about the strength of our confidence in

[35]Some have suggested that degrees of belief are just explicitly probabilistic categorical beliefs, e.g., a belief that the probability of rain today is 50%. For example, Harman (1986) suggests that in the few cases we seem to have degrees of belief, they are actually just explicitly probabilistic categorical beliefs. I'll say more about the question of reducing degrees of belief to categorical belief. For an excellent discussion of this issue, see Christensen (2004, pp. 18-20).

[36] You might even think that the traditional, categorical doxastic attitudes are different grades of belief in a doxastic hierarchy with only three levels of strength: belief, suspension, and disbelief.

different claims, and about changes in the strength of our confidence in particular claims in light of changes in our evidence.

What gives rise to the gradational structure of confidence? Do we have different degrees of belief somehow in isolation, or do we have different degrees of belief only holistically in virtue of differences between our degrees of belief in different claims? We often talk about isolated degrees of belief, saying things like “I’m pretty sure that Biden is French” or “I’m certain that Trump won’t get the nomination.” However, it’s hard to imagine how degrees of belief could exist in isolation. Rather, it seems that when we talk about individual degrees of belief in isolation, we are implicitly placing them in a kind of confidence ordering. When one says “I’m pretty sure that Biden is French,” one implicitly places *Biden is French* near, but not quite at the top of the ordering.

Assuming that the gradational structure of degrees of belief does arise out of confidence comparisons, we confront the further question of whether all claims are comparable in terms of confidence. That is, does each agent have a univocal confidence ordering in which they compare all claims? Or are some claims incomparable in terms of confidence for some of agents, resulting in multiple incommensurable confidence orderings? The typical probabilist representation of degrees of belief with real-valued functions suggests that all of an agent’s degrees of belief are comparable in a univocal confidence ordering. But, intuitively, it seems that some claims are incomparable—or at least very hard to compare with much precision. For example, could you precisely compare the following two claims: *the population of Tokyo will be 40 million by 2025*, and *the first card drawn from my old and likely incomplete deck will be a spade*?[37]

For the most part, probabilists don't give explicit answers to these questions. However, in some cases, they do seem to commit themselves to having specific answers—for example, their formal representation of degrees of belief and their arguments for the probabilist norm seems to provide such commitments. I'll call attention to some of these implicit commitments below.

### 2.1.4 Degrees of Belief & Other Doxastic Attitudes

Epistemologists generally acknowledge two main kinds of doxastic attitudes: degrees of belief and the categorical attitudes I've already mentioned. Some epistemologists also acknowledge a third kind of attitude called "comparative confidence." Comparative confidence is an attitude an agent bears to pairs of claims. An agent's comparative confidence in a pair of claims is a comparison of the strength of the agent's confidence between claims. The agent might be strictly more confident that $A$ than that $B$, or *vice versa*. Or the agent might be equally confident in $A$ and $B$. Or their confidence in $A$ and $B$ might not be comparable at all. Comparative confidence can be construed as an alternative conception of graded doxastic attitudes, as one's comparative confidence assessments form a kind of confidence ordering that can be thought to convey information about one's degrees of confidence.

[37] This example is an adaptation of one from Fishburn (1986, p. 339), which was in turn inspired by Keynes (1921).

When probabilists present the first tenet of probabilism (if they do so explicitly at all), they do not usually explain how degrees of belief relate to these other kinds of doxastic attitudes. Probabilists have three main options for specifying the connections between degrees of belief, the categorical attitudes, and comparative confidence: pluralism, reductionism, and eliminativism.[38] Pluralism is the view that categorical belief,[39] degree of belief, and comparative confidence are all legitimate, distinct doxastic states. None of these states reduce to each other; none of these notions should be eliminated from philosophical discourse. Reductionism is the view that one or two of the states reduce to one of the others. Eliminativism is the view that one or more of the notions should be removed from philosophical discourse. The eliminable notions are empty or useless; they do not contribute to philosophical discourse in any important way.

*Prima facie*, probabilists can endorse any of these options so long as they hold that at least some doxastic states—degrees of belief—are graded. So, it looks like the only view unavailable to probabilists is eliminativism about *degree of belief*. Pluralism seems to be a live option. So do the various forms of reductionism. Typically, however, the probabilists who have countenanced the issue of the relationship between these

[38] For a nice survey of these options, see Christensen (2004, Ch. 2). Christensen focuses on the connections between degree of belief and categorical belief, but his discussion can easily be adapted to apply to comparative confidence too. Hawthorne (2009) is, to my knowledge, the only discussion of the connections between all three kinds of attitudes.

[39] Rather than listing all of the categorical doxastic attitudes, I'll focus on categorical belief hereafter.

notions have focused on the relationship between categorical belief and degree of belief, and they have favored reductionism and eliminativism of various kinds. Several probabilists have argued that the notion of categorical belief should be replaced with degree of belief in our epistemologies.[40] Others have argued that categorical belief reduces to degree of belief.[41] Few authors have invoked the notion of comparative confidence in this discussion.[42] Thus, despite the variety of options for interpreting and further characterizing the first tenet of probabilism, many probabilists present it as the claim that beliefs come in degrees—suggesting the eliminativist or reductionist reading.

### 2.1.5 Formal Representations

On its own, the first tenet of probabilism doesn't seem to say anything about whether—or how—degrees of belief should be represented formally. It merely posits a graded doxastic state. But the second tenet of probabilism says that degrees of belief ought to conform to the axioms of the probability calculus. And since, strictly speaking, the probability axioms apply to real-valued functions, it seems that probabilism requires some kind of formal representation of degrees of belief in terms of real-valued functions.

---

[40] Jeffrey (1970) is the most well-known proponent of this position.

[41] See, for example, Sturgeon (2008), Foley (2009), and Hawthorne (2009).

[42] For some notable exceptions see Koopman (1940) and Hawthorne (2009). While few authors have written about how comparative confidence relates to degree of belief and categorical belief, there is a fairly long tradition of work on comparative confidence and comparative probability among Bayesians. See Fine (1973), Fishburn (1986), and Capotorti & Vantaggi (2000) for overviews of some of this work.

So, probabilists usually represent an agent's degrees of belief formally in terms of a pair of credence functions, an unconditional credence function, $cr(\cdot)$, and a conditional credence function, $cr(\cdot \mid \cdot)$, each of which assigns a number, $r \in [0,1]$, to the expressions of a formal language.[43]

Credence function values closer to 1 represent higher degrees of confidence; values closer to zero represent lower degrees of confidence; .5 represents equivocal confidence. ⌜ $cr(A)$ ⌝ represents the agent's degree of belief in the claim represented by ⌜ $A$ ⌝. ⌜ $cr(A|B)$ ⌝ represents the agent's degree of belief in the claim represented by ⌜ $A$ ⌝ on the assumption that the claim represented by ⌜ $B$ ⌝ is true. So, for example, ⌜ $cr(A)$=1 ⌝ represents that the modeled agent is certain in the truth of $A$, ⌜ $cr(A|\neg B) = 0$ ⌝ represents that the agent is certain that $A$ is false on the assumption that $B$ is false, and ⌜ $cr(B) = .5$ ⌝ represents that the agent is equally confident that $B$ is true as that it is false.

This formal representation of degrees of belief by means of such a pair of real-valued credence functions captures the graded nature of degrees of belief. It also permits a connection between degrees of belief and the probability axioms. A full account of the probabilist scheme for formally representing degrees of belief should include more explicit and precise connections between degrees of belief and their

[43] Probabilists disagree about which formal language to use to represent the objects of degree of belief. Some prefer the language of set-theory; others prefer languages for sentential or predicate logic. For my purposes in this section, the differences don't really matter, so I present a formal representation in terms of the sentences of a logical language. I'll say more about this issue in §2.3.1 and Chapter 4.

numerical representations. In particular, it should include an account of the precise significance of the representation of degrees of belief with real numbers, *per* the discussion in §2.1.4. Although probabilists do not usually present the details of this sort of account,[44] I'll present some of these detail in the next two sections.

Also, as I mention above, probabilists are typically reductionists or eliminativists about the notions of categorical belief and comparative confidence. So, this formal representation is often meant to be an exhaustive account of the doxastic states of epistemological interest. However, one might propose a formal representation for the states of categorical belief and comparative confidence as well. But, since most probabilists treat credence functions as sufficient for formally representing all doxastic states, their treatments suggest that they take the probability axioms to be sufficient norms for all doxastic states.

## 2.2 Why Posit Degrees of Belief?

Probabilists posit degrees of belief for two main kinds of reasons: to better describe our doxastic states in order to better account for our experiences and behavior; and to provide a more plausible normative account of what our doxastic states should be like.[45] I've said quite a lot so far about what degrees of belief are, and I've hinted at some of the reasons for positing them, but now I will present these reasons more explicitly, in more detail.

---

[44]Titelbaum (2013) is a notable exception.

[45] I follow Titelbaum (MS) in distinguishing descriptive and normative reasons for positing degrees of belief. My discussion of the normative reasons for positing degrees of belief follows Titelbaum's presentation quite closely.

### 2.2.1 Descriptive Reasons

Probabilists cite three main descriptive reasons for positing degrees of belief. First, introspection and our practices of mental state attribution (including self-attribution) reveal that we often have different degrees of confidence in various claims.[46] We may find ourselves pretty confident that it won't rain, and even more confident it won't snow. We also say things like "I'm more confident that Clinton will win than that Walker will win," and "I'm very confident that Trump won't get the nomination." Sometimes we even quantify the strength of our confidence: "I'm 95% sure I turned off the stove." These cases suggest that there is a graded doxastic state that supports pretty fine-grained comparisons of relative strength. Probabilists posit degrees of belief to describe these cases, rather than try to account for them in terms of the traditional, categorical doxastic states of belief, disbelief, and suspension of belief.

Second, probabilists posit degrees of belief to account for the way we respond to evidence.[47] We generally try to match our doxastic attitudes to our evidence. But, probabilists note, our evidence can be weaker or stronger, and it often changes incrementally over time. For example, imagine that I believe my spouse is at home right now on the basis that this would conform to her usual schedule and she told me her plans to work from home when I left for the office this morning. When I see that she is posting pictures of our cats (who reside only at home) to social media, my confidence

[46] See, for example, Christensen (2004, p. 13), Huber (2009, p. 1), Easwaran (2011, p. 312), and Titelbaum (MS, Ch. 1).

[47] See, for example, Christensen (2004, p. 13), Joyce (2004, p. 133), and Titelbaum (MS, Ch. 1).

that she is home increases further. Even though I already have a categorical belief that my spouse is at home, my confidence that she is at home grows as the evidence mounts. We often experience these kinds of shifts in our confidence even when our categorical beliefs stay the same. On the assumption that our doxastic attitudes can match the evidence fairly closely, cases like this suggest, once again, that there is kind of a graded doxastic state, which can be weaker or stronger, and which can change incrementally over time to match the evidence.

Third, probabilists posit degrees of belief to better describe and explain our behavior and decision-making.[48] Betting behavior, in particular, is one of the phenomena degrees of belief are most commonly invoked to explain. My offer of 4:1 odds for Stewball winning the next race indicates that I am less than certain that Stewball will win, but much more confident than not that he will.[49] Accounting for my behavior in this case seems to require degrees of belief.

However, explicit betting behavior is not the only behavioral source of evidence for degrees of belief. There are many more commonplace examples. Imagine, on a rainy day, that I opt for my golf umbrella rather than my raincoat, despite the fact that I believe both will keep me dry and I have no other preference for one over the other. The best explanation for my choice in this case is that, while I believe either option will keep me dry, I'm more confident that the umbrella will do so than that the raincoat will.

---

[48] See, for example, Christensen (2004, pp. 13-14), Joyce (2004, p. 133), and Titelbaum (MS, Ch. 1).

[49] This example is borrowed from Joyce (2004, p. 133).

Similarly, imagine that I enjoy accountancy and marine biology equally, and I believe I could get a job pursuing either career, but I choose to pursue a career as an accountant rather than as a marine biologist. One plausible explanation of my choice in this case is that I'm more confident that I'll get a job as an accountant. My behavior in all such cases seems to demand explanation in terms of a doxastic state that is more finely grained than the categorical states alone.

On their own, these cases give probabilists pretty strong reason to posit degrees of belief. However, many probabilists think the case for degrees of belief is further buttressed by the lack of plausible alternative explanations of the phenomena described above.

Attempts to explain away the appearance of degrees of belief in our introspection and attitude reports don't seem to have any hope of getting off the ground. So, it doesn't look like we can avoid positing degrees of belief by denying or reducing the phenomena they're posited to describe and explain.

The most commonly proposed way of explaining away degrees of belief is to account for their appearance in terms of categorical beliefs about probabilities.[50] On this approach, when I say "I'm very confident Trump won't get the nomination," I'm not expressing a graded doxastic state with a strong negative valence towards Trump's candidacy. Instead, I'm expressing a categorical belief that it is very improbable that Trump will get the nomination. Similarly, on this view, my response to seeing my

[50] This is the kind of response probabilists often encounter at conferences and in the classroom. Its most cited occurrence in print is due to Harman (1986).

spouse's posts of cat pictures on social media is not an increase in the strength of my degree of belief that she is at home. Rather, this evidence causes me to change my categorical belief about my estimation of the probabilities that she is at home. And my selection of the umbrella over the raincoat is not explained by my higher degree of confidence that the umbrella will keep me dry. Rather, it is explained by my categorical belief that it's more probable that umbrella will keep me dry than that the raincoat will. Since we can explain away the cases that motivate positing degrees of belief, we don't need to posit them, or we can reduce talk about them to talk about categorical beliefs. Degrees of belief are theoretical dead weight.

Probabilists have identified several problems with this approach. So, they think it's more natural to simply posit degrees of belief rather than try to explain them away in this way. It seems like we can have degrees of belief without the attendant categorical beliefs about probabilities.[51] For, one may have degrees of belief despite being ignorant of the concept of probability. Perhaps some children and non-human animals are like this, for example.[52] This suggests that to reduce degrees of belief to categorical beliefs about probabilities is to assume an implausible degree of probabilistic sophistication in agents. Or one might be probabilistically sophisticated, and have a degrees of belief about, say, whether Trump will get the Republican nomination, but not yet have formed

---

[51] Eriksson & Hájek (2007, pp. 206-207) provide a few such cases, similar to those I offer in this paragraph.

[52] See Frankish (2009, p. 77).

a categorical belief explicitly about the probability that Trump will get the nomination.[53] This kind of case suggests that the reduction of degrees of belief to categorical beliefs would make categorical beliefs about probabilities more common than they actually are. Alternatively, one might be very probabilistically sophisticated, while still recognizing in oneself that one's degrees of belief diverge from one's probability estimates. For example, I might find that, while I estimate the logical probability, evidential probability, and metaphysical chance (propensity) of Trump getting the nomination to be very low, I have a strong and irrepressible degree of belief that he'll get the nomination. In such a case, I might be motivated to revise my degrees of belief or my probability estimates. Indeed, we might even view it as a good thing to try and match our degrees of belief to our beliefs about probabilities of various kinds (for example, to our beliefs about evidential probabilities, or to our beliefs about objective chances).[54] But such a view assumes a distinction between degrees of belief and beliefs about probabilities. Because of cases like this, probabilists are skeptical that degrees of belief can be explained away in terms of categorical beliefs.

In addition to pointing out cases where degrees of belief and categorical beliefs about probabilities seem to diverge, probabilists also find other flaws with attempts to reduce degrees of belief to categorical beliefs about probabilities. Christensen (2004, p. 18-20), for example, raises the issue of what kind of probability it is that degrees of belief are being reduced to beliefs about. For philosophers, scientists, and mathematicians

[53] Titelbaum (MS) describes cases like this.

[54] That's the idea behind the principal principle.

have distinguished several different kinds of probability: relative frequency, logical probability, evidential probability, metaphysical chance, and subjective probability.

The kind of probability involved in such a reduction has a significant impact on the plausibility of any such reduction. For example, if we go for a subjective interpretation of probability, the reduction seems to collapse. On that view, degrees of belief would be reduced to beliefs about subjective probabilities, but degrees of belief and subjective probabilities are usually thought to be one and the same.[55] So, on such an account degrees of belief would turn out to be beliefs about degrees of belief.

If we choose any of the other options (logical probability, evidential probability, metaphysical chance, etc.), we run into sophistication issues and counterexamples. For, it is doubtful that a very large portion of the population has—with much regularity—categorical beliefs about relative frequencies, evidential probabilities, logical probabilities, or metaphysical chances *as such*. Indeed, even well-informed people may explicitly doubt the existence of logical probabilities and objective chances, and may doubt that relative frequencies are relevant to single cases—for example, that Trump will win the nomination. But they may still have various degrees of confidence in such claims. Furthermore, even if there is a way to make sense of degrees of belief as beliefs about probabilities in a way that accounts for the ubiquity with which we seem to

[55] To call subjective probabilities "probabilities" suggests that they conform to the probability axioms, which most of our degrees of belief likely do not. Nonetheless, in common probabilist parlance, "degree of belief" and "subjective probability" are often used interchangeably. I should note, however, that there is another sense of "subjective probabilities," which refers to the degrees of belief of ideal agents, the kind of agents whose degrees of belief, the story goes, would satisfy the probability axioms.

recognize degrees of belief in ourselves and others, we would still need to account for cases in which probabilistically sophisticated individuals have degrees of belief that diverge from their estimates of relative frequencies, logical and evidential probabilities, and metaphysical chances.

So, while accounting for degrees of belief in terms of categorical beliefs about probabilities has some initial intuitive appeal, in light of the issues I've just described, probabilists hold that it's more plausible to simply posit degrees of belief as a kind of mental state in their own right. As I mention in §2.1.4, some probabilists attempt to reduce categorical beliefs to degrees of belief. Others favor eliminating categorical beliefs all-together. Still others acknowledge both kinds of doxastic state and attempt to characterize the normative and descriptive connections between them without proposing any reduction of one to the other.

**2.2.2 Normative Reasons**

The dominant approach to epistemology focuses on categorical belief. It draws on several widely accepted norms for categorical belief, most prominently, the consistency and closure norms. According to the consistency norm, one's categorical beliefs ought to be deductively consistent. So, for example, one shouldn't believe that Clinton will be the next President and that Bush will be the next President. According to the closure norm, our beliefs ought to be closed under deductive entailment. So, if two or more of your beliefs jointly entail another claim, then you should believe that claim too.[56] For

[56] Less demanding forms of the closure norm might not require that you actually believe all of the logical consequences of your beliefs. Instead, it might require that you should

example, if you believe that if Bush wins, then Clinton will not, and you believe that Clinton will win, you should also believe that Bush will lose.

The consistency and closure norms are pretty intuitively appealing. They express the commonly held view that logic provides norms for belief, and they generate plausible normative verdicts in many cases, like the examples immediately above.

In other cases, however, the verdicts generated by the consistency and closure norms are less plausible. In general, the consistency and closure norms do not stand up to very close scrutiny—at least not as I have stated them above. Both norms seem to be too demanding. Unrestricted, the consistency norm demands that we maintain the consistency of all of our beliefs—a huge number—even in the face of misleading or conflicting evidence (as is often the case). The closure norm demands that we believe all of the logical consequences of our beliefs—including even the most complex tautologies. Furthermore, in some cases, the consistency and closure norms are not only too demanding; they also seem to make the wrong normative demands. For example, the closure norm tells me that, if I believe that Joe Biden is a Republican or an octopus from outer space, and I believe that Biden is not a Republican, then I should believe the he is a space octopus. That is, the closure norm says nothing of the fact that sometimes we should revise our beliefs rather than adopt their logical consequences. Similarly, there may be cases in which we should maintain inconsistent beliefs rather than

be merely disposed to believe these consequences if you should have the occasion to entertain them.

disregard our evidence.[57] In addition to these general problems with the consistency and closure norms, probabilists focus on two particular counterexamples to these norms, the lottery paradox and the preface paradox.

The lottery paradox goes like this.[58] Imagine a fair lottery with one million tickets. Imagine further an agent who has purchased a ticket in this lottery. Because the lottery is large and fair, the agent believes that their ticket will not win, and furthermore they believe of each other ticket that it will not win. But because the lottery is fair, the agent believes that some ticket will win.

The paradox here is that the agent's beliefs seem rational, but they violate the intuitively plausible consistency and closure norms. The probability of any one ticket winning is vanishingly small, and no ticket is any more likely to win than any other. So, the agent's belief of each ticket that it will not win seems reasonable enough. And, because the lottery is fair, it makes sense for the agent to believe that some ticket will win. The problem is that these beliefs are inconsistent—they can't both be true at once. Thus, the lottery paradox seems to be a counterexample to the consistency norm. It also seems to provide a counterexample to closure: the beliefs, for each ticket, that *this ticket won't win* jointly entail (under closure) the belief that no ticket will win.

---

[57] See, for example, Harman (1984, pp. 108-109). The lottery and preface cases may be just such cases.

[58] The lottery paradox is due to Kyburg (1961).

The preface paradox goes like this.[59] Imagine an agent who has written a long, carefully argued work of non-fiction. The main text of this book includes many claims, all of which the agent believes. Imagine further that like many other non-fiction authors, the author includes in the preface the claim that there are bound to be some mistaken claims in the sequel, for which the author alone—not their many helpful colleagues—bears full responsibility. Imagine that the author is not just being modest; they really do believe that at least some mistakes—some false claims—have creeped into the main text.[60] Are the author's beliefs incoherent?

Like in the lottery case, the agent's beliefs seem intuitively rational. The author has carefully crafted the claims in the main text, adduced evidence toward them and considered objections. Belief in these claims seems rational. Yet it also seems reasonable to believe that a long book like the author's will include at least one mistaken claim. Once again, the author's beliefs are jointly inconsistent and violate closure (unless the author believes every claim). If the preface belief is true, then some of the beliefs asserted in the main text are false, and if the beliefs in the main text are true, then the preface belief is false. The author can't consistently hold them all while maintaining closure.

---

[59] The preface paradox is due to Makinson (1965).

[60] If you find the preface story implausible, consider a structurally similar case. Consider your own belief set. Trivially, you believe all of your beliefs. But you likely also believe that at least some of your beliefs are mistaken. This case too violates consistency and closure. And it hits very close to home. Easwaran & Fitelson provide such a global version of the preface paradox in their (2015).

In light of the lottery and preface paradoxes, and the general concerns about demandingness and normative appropriateness mentioned above, things don't look good for the consistency and closure norms for categorical belief. However, if we posit degrees of belief and probabilistic norms for degrees of belief, and if we re-consider the lottery and preface cases in terms of degrees of belief and probabilistic norms for them, things don't look so bad.

Re-imagined thus, the lottery case is one where the agent has very low degrees of belief that each ticket will win, and a very high degree of belief that some ticket will win. Construed this way, the paradox dissolves. It seems perfectly rational to be very confident that each ticket of a fair lottery will lose, while at the same time being very confident that some ticket will win. Indeed, it seems that these are precisely the degrees of belief one ought to have in this case.

Similarly, a re-imagination of the preface case in terms of degrees of belief goes like this. The author has high degree of belief that each of the claims of the main text is true, but the author also has low degree of belief that they're all true. Again, these degrees of belief seem eminently reasonable.

In both degree of belief based re-imaginations, the intuitive verdicts that the lottery and preface degrees of belief are rational is supported by a probabilistic coherence norm for degrees of belief analogous to the deductive consistency norm for categorical beliefs. I'll present the probabilist norm in more detail in §2.3. Thus, probabilists posit degrees of belief, in part, because doing so allows them to provide more plausible normative verdicts in cases like the lottery and the preface.

Like the descriptive reasons for positing degrees of belief presented above, probabilists' normative reasons have also been criticized. Some critics have argued that more careful statements of the consistency and closure norms will make them less demanding and more plausible. Others have criticized the lottery and preface cases, arguing that the agents' beliefs in these case are not rational, so they don't serve as counterexamples to the consistency and closure norms. Others try to avoid positing degrees of belief and probabilistic degree of belief norms by proposing alternative (sometimes probabilistic) norms for categorical belief. Still others argue that positing degrees of belief and probabilistic degree of belief norms fails to provide genuine solutions to the lottery and preface paradoxes, which are paradoxes about categorical belief, not degree of belief. Degree of belief responses just change the subject.

Ultimately, probabilists admit that, on their own, none of the reasons presented in this section give knock-down support for positing degrees of belief. Rather, they contend that together these reasons make it more natural to posit degrees of belief than to try do without them, and they note that once we do posit degrees of belief and invoke the machinery of subjective Bayesianism, we reap significant theoretical rewards. In this dissertation, I will take up some of the issues surrounding the reasons presented in this section. In particular, I will cast doubt on the view that probabilistic norms for degrees of belief are much less problematic than the consistency and closure norms for categorical belief. However, from here on out I will assume that there is a graded doxastic state that can be weaker or stronger by degrees, and I will call this state "degree of belief."

### 2.3 What Is the Probabilist Norm?

According to the second tenet of probabilism, our degrees of belief ought to conform to the axioms of the probability calculus. As I note above, a naïve interpretation of this claim sounds like a category mistake. Degrees of belief are doxastic states, which might be understood to be mental representations, mental or behavioral dispositions, etc. The probability axioms characterize real-valued functions. *Prima facie*, it's a bit strange to say that these axioms for real-valued functions are norms for doxastic states. But most probabilists wouldn't bat an eye at this claim. Indeed, this is exactly how probabilists usually talk about the probability axioms. There are two apparent reasons for this. First, probabilists represent degrees of belief formally in terms of functions from sentences to numbers. And, second, probabilists tend to be a bit sloppy about distinguishing their formal modeling frameworks from the things those frameworks are supposed to model. In this section, I'll present the probability calculus and the typical probabilist interpretation of its axioms and theorems as degree of belief norms.

#### 2.3.1 The Probability Calculus

Several formal theories of probability—several probability calculi—have been proposed throughout the development of mathematical probability theory. However, when probabilists talk about *the* probability calculus, they usually have in mind Kolmogorov's (1956) axiomatization, which is widely accepted by mathematicians, scientists, and philosophers.[61] In this subsection, I'll present Kolmogorov's probability calculus.

Kolmogorov stated his probability calculus in set-theoretic terms, and this set-theoretic version of his calculus is typical in the philosophical literature on probability. However, a version of Kolmogorov's calculus stated in logical rather than set-theoretic terms is also common. In Chapter 4, I will argue that the logical approach is preferable to the set-theoretic approach in the context of probabilism. So, here I will present both versions. I'll present the set-theoretic approach first, and then I will briefly present the logical approach and describe some of the important differences between the two approaches.

Before I get into the details of Kolmogorov's probability calculus, I want to emphasize that it can be viewed on its own as a piece of pure mathematics, without regard to any particular interpretation. In this sense, the probability calculus is a characterization of a particular kind of mathematical structure: it characterizes a set and a pair of functions from that set (or its Cartesian product) to the set of real numbers. This structure is called a "probability model" and the functions are called "probability functions" because they are supposed to capture many of the essential features of our

[61] While Kolmogorov's calculus is widely accepted, it is not without its critics, and some alternative formal theories of probability and uncertainty have been proposed. For excellent surveys of criticisms of Kolmogorov's calculus, see Fine (1973, especially Ch. 3) and Lyon (MS). For a survey of alternative formal theories of probability and uncertainty, see Huber (2014). The most well-known alternative formal theories of probability are due to Popper (1955) and Rényi (1955, 1970). These two alternatives provide generalizations of Kolmogorov's axioms; Kolmogorovian probability functions are special cases of Popper functions and Rényi functions.

everyday concept of probability. However, probability models admit other interpretations as well. In this dissertation, I will focus on the use of the probability calculus as a logic for degrees of belief. So, in this section, although I will present Kolmogorov's calculus primarily as a piece of mathematics, I will also describe its intended interpretation as a model of the everyday notion of probability along the way. Then, in §2.3.2, I will show how probabilists apply the probability calculus to generate norms for degrees of belief.

On the set-theoretic approach, a probability model or "probability space" consists of a non-empty set, $\Omega$, a field, $\mathcal{F}$, of subsets of $\Omega$, and a unary probability function, $pr(\cdot): \mathcal{F} \longrightarrow \Re$.

The set, $\Omega$, which is often called the "sample space," could be any non-empty set, but it is usually taken to be a set of so-called "elementary events," the set of all possibilities or possible worlds relevant to an application.

The field, $\mathcal{F}$, is a set of subsets of $\Omega$, which includes $\Omega$, and which is closed under complementation (with respect to $\Omega$) and finite intersections of members of $\mathcal{F}$. So, if $\mathrm{A}$ is a subset of $\Omega$ in $\mathcal{F}$, its complement, $\overline{\mathrm{A}}$ (that is, $\Omega - \mathrm{A}$, the set of all members of $\Omega$ not in $\mathrm{A}$), is also in $\mathcal{F}$; and if $\mathrm{A}$ and $\mathrm{B}$ are subsets of $\Omega$ in $\mathcal{F}$, then their intersection, $A \cap B$, is also in $\mathcal{F}$.[62] One example of a field of subsets of $\Omega$ is its power set, the set of all of its subsets (Easwaran 2011, p. 319, n. 1).

[62]Often when probability theorists characterize $\mathcal{F}$, they say it is a set of subsets of $\Omega$ closed under complementation and finite unions. However, this is a difference that doesn't make a difference. If $\mathcal{F}$ is closed under finite intersections and complementation, then it is also closed under finite unions.

The probability function, $pr(\cdot)$, is a one-place function that maps the members of $\mathcal{F}$ to members of the set of real numbers according to the following axioms.

For all sets $\mathrm{A}, \mathrm{B} \in \mathcal{F}$:

1. $pr(\mathrm{A}) \geq 0$; (non-negativity)
2. $pr(\Omega) = 1$; and (normality)
3. If $\mathrm{A} \cap \mathrm{B} = \emptyset$, then $pr(\mathrm{A} \cup \mathrm{B}) = pr(\mathrm{A}) + pr(\mathrm{B})$. (finite additivity)

These axioms guarantee that probability functions are non-negative, normalized (with a highest value of 1), and finitely additive. Axiom 1 guarantees that probability values are non-negative. Every output of a probability function must be greater than or equal to zero. It also guarantees that probability functions are total functions: there is a real number output for every element of $\mathcal{F}$ input. Axiom 2 normalizes probabilities with a highest value of 1. Where $\Omega$ is taken to be a set of events, possibilities, or propositions, axiom 2 is a requirement that necessities or logical truths (depending on the kind of possibilities etc. we're dealing with) receive probability 1. Axiom 3 says that probabilities of disjoint members of $\mathcal{F}$ are finitely additive.

In addition to these three axioms for $pr(\cdot)$, Kolmogorov's probability calculus also includes a definition of a binary conditional probability function, $pr(\cdot \mid \cdot): \mathcal{F} \times \mathcal{F} \longrightarrow \Re$ with the following formula. For all $\mathrm{A}, \mathrm{B} \in \mathcal{F}$:

4. If $pr(B) \neq 0$, then $pr(\mathrm{A}|\mathrm{B}) = \frac{pr(\mathrm{A} \cap \mathrm{B})}{pr(\mathrm{B})}$. (ratio formula)

This binary conditional probability function, $pr(\cdot \mid \cdot)$, is intended to capture the probability of an event (proposition, possibility, etc.) given—or *on the assumption* of—the occurrence (truth) of another event (proposition, possibility, etc.). The ratio formula defines conditional probabilities in terms of a ratio of unconditional probabilities. I want to emphasize that the ratio formula gives a *mathematical* definition of conditional probability functions in terms of unconditional probability functions. We should not assume without argument that this mathematical definition carries over to the concept of probability it is intended to capture. Several philosophers have rejected such a definition, and some have argued that we should treat conditional probability functions as basic in our formalizations of probability.[63] I will not pursue the adequacy of the ratio formula as an analysis of conditional probability in this dissertation. I will, however, discuss its role in providing norms for degrees of belief in the next sub-section.

Axioms 1-3 and the ratio formula make up the core of what is usually referred to as "the probability calculus." As I note above, these formulas are mathematical abstractions, and as such they are not terribly intuitive without interpretation. Furthermore, they are chosen because they provide a concise and relatively uncontroversial account of the properties that probabilities must have in order for them to have various intuitive and desirable properties. They don't themselves state properties that laypeople typically associate with the everyday concept of probability. But together—with a little set theory and algebra—they have many easily derived consequences that capture features that fit more intuitively with the everyday concept of probability.

[63] See, for example, Hájek (2003) and Hawthorne (2011).

In addition to the core axioms and definition of the probability calculus, Kolmogorov also proposed an additional constraint on probabilities known as the "countable additivity" axiom. This axiom applies in place of the finite additivity axiom on the assumption that $\mathcal{F}$ is a $\sigma$-field—that is, a set of subsets of $\Omega$ closed under complementation and *countable*—rather than finite—intersections. Assuming $\mathcal{F}$ is a $\sigma$-field, $pr(\cdot)$ is a countably additive "probability measure" if it satisfies axioms 1 & 2 and the following additional axiom.

5. If $\{\mathrm{A}_i\}$ is a countably infinite set of mutually disjoints sets (that is, for distinct $\mathrm{A}_i$ and $\mathrm{A}_j$ in $\{\mathrm{A}_i\}$, $\mathrm{A}_i \cap \mathrm{A}_j = \emptyset$), then $pr(\cup\{\mathrm{A}_i\}) = \sum_i pr(\mathrm{A}_i)$.[64] (countable additivity)

For various reasons, the countable additivity axiom is controversial.[65] So, in this dissertation, I will consider it an additional constraint separate from the probability calculus. I will not pursue this issue further in this dissertation.

So, that's the set-theoretic version of Kolmogorov's probability calculus. On the logical approach, we substitute the machinery of set theory for that of classical deductive logic. In stating the probability calculus, this means that we replace $\Omega$ and $\mathcal{F}$ of the set-theoretic approach with a formal logical language, $\mathcal{L}_n$, with $n$ atomic sentences closed under the usual logical operations.[66] So, on the logical approach, a

[64] Weisberg (2011, p. 504).

[65] For a critical discussion, see Howson & Urbach (2006, pp. 26-29).

[66] $\mathcal{L}_n$ could be a sentential or predicate language. If $\mathcal{L}_n$ is a finite sentential language, then the probability calculus is decidable. If $\mathcal{L}_n$ is a first order language for predicate logic, we gain further expressive power, but at the cost of decidability.

unary function $pr(\cdot): \mathcal{L}_n \rightarrow \Re$, is a probability function if and only if it satisfies the following axioms, where the complement, union and intersection are replaced with the logical operations of negation, conjunction, and disjunction, and the sample space is replaced with the notion of tautology. Read $\ulcorner \vDash A \urcorner$ to say that $A$ is a tautology.

For all sentences $A, B \in \mathcal{L}_n$:

1. $pr(A) \geq 0$; (non-negativity)
2. If $\vDash A$, then $pr(A) = 1$; and (normality)
3. If $\vDash \neg(A \wedge B)$, then $pr(A \vee B) = pr(A) + pr(B)$. (finite additivity)

Carrying on with the substitution, we define a binary conditional probability function, $pr(\cdot \mid \cdot): \mathcal{L}_n \times \mathcal{L}_n \longrightarrow \Re$ with the following formula. For all $A, B \in \mathcal{L}_n$:

4. If $pr(B) \neq 0$, then $pr(A|B) = \frac{pr(A \wedge B)}{pr(B)}$. (ratio formula)

For many purposes, the differences between the set-theoretic and logical presentations are mainly superficial—just a matter exchanging some set-theoretic notation for logical notation. When the probability calculus is considered as a formal abstraction, the set-theoretic approach may be preferable in the sense that set theory is a more general formal framework. For under a typical interpretation, the logical approach induces the same set-theoretic structure as that required by the set-theoretic approach (Huber 2014, §1.3). However, set theory itself presupposes a logic in which it is couched—predicate logic. In any case, in the context of normative modeling for degrees of belief, there are points in favor of both approaches. In Chapter 4, I will argue at length that the logical approach is preferable for the purposes of probabilism, so that is the approach I will assume until then.

### 2.3.2 Treating the Probability Axioms as Norms

Since probabilists represent degrees of belief with real-valued functions, the most obvious interpretation of the second tenet of probabilism is that the credence functions that represent an agent's degrees of belief should be probability functions. Thus, the credence functions that represent an agent's degrees of belief should be non-negative, normalized, and finitely additive. And an agent's conditional degrees of belief should afford representation as probabilities *per* the definition of conditional probability. But, additionally, the second tenet of probabilism suggests that the axioms (and theorems) of probability theory should serve as norms for the apportionment of degree of belief.

It's not obvious how this is supposed to work. But a simple answer is that the probability axioms (and theorems) should be translated into doxastic norms by appeal to the typical probabilist scheme for representing degrees of belief in terms of degree of belief functions. The idea is that we can use this representation scheme in reverse to translate norms about real-valued functions into norms about degrees of belief.[67] This seems to be how most probabilists interpret the claim that degrees of belief should obey the probability axioms.

Unfortunately, not all of the probability axioms themselves afford straightforward translation into norms for the apportionment of degrees of belief. But, given the nice features and implications of the probability axioms described above, we get some pretty intuitive norms when we consider the probability calculus, properly interpreted, as a

---

[67]The term "representation scheme" is adopted from Titelbaum (2013).

whole. For example, it's not clear what the first axiom, which requires probabilities to be non-negative, means for degree of belief. Considered on its own, in light of the typical probabilist representation scheme, it says something like an agent should have at least some degree of confidence in every claim. Thus, agents ought to opinionated. But in conjunction with the other axioms, axiom 1 implies that one should be sure of the falsehood of all contradictions, and one cannot be less confident in any claim than one is in a contradiction. Considered on its own, the second axiom says that one should be certain of all logical truths. And, in conjunction with the other axioms, the result is a requirement of logical omniscience. One should be sure of all tautologies, sure that all contradictions are false, and sure of all entailment relations. And it says that one can never be more confident in a claim than one is in a tautology. The third axioms says that, if two claims are logically incompatible, one's confidence in the disjunction of the two should equal the sum of the two. The ratio formula says that the degree of one's confidence in $A$ given $B$ should be equal to the ratio of the degree of one's confidence that $A$-and-$B$ compared to the degree of one's confidence in $B$. So, for example, if you are very confident that the game will be canceled on the assumption that it will rain, your confidence that it will rain and the game will be canceled should be nearly as strong as your confidence that it will rain.

Here are two of the more intuitive theorems of the probability calculus.

Entailment theorem   If $A \vDash B$, then $pr(B) \geq pr(A)$, for all $A, B$.

Negation theorem   $pr(\neg A) = 1 - pr(A)$, for all $A$.

Translated into a degree of belief norm, the entailment theorem says something like the following:

Entailment norm — If $A \vDash B$, then you should be at least as confident in $B$ as you are in $A$.

And when we translate the negation theorem into a norm we get something like this:

Negation norm — Your confidence in $\neg A$ should be as strong as your confidence in $A$ is weak, and *vice versa*.

Both of these norms seem like eminently reasonable constraints on degrees of belief.

Together the probability axioms and theorems, interpreted by means of the representation scheme for degree of belief, impose a kind of coherence requirement on degree of belief. They also require logical omniscience and opinionation. These are very strong norms for degree of belief apportionment. For this reason, many philosophers have objected to probabilism for providing implausible, overly-idealized norms. But, as I suggest above, this account of the formal representation of degrees of belief and the normative role of the probability calculus is not the only interpretation available to probabilists. In the next chapter, I will present some of the common objections to probabilism. And in later chapters, I will show how a more plausible formal representation and account of the normativity of probability can avoid these objections.

### 2.4 Justifications for the Probabilist Norm

Probabilists give three main kinds of arguments to justify the probabilist norm: Dutch book arguments, representation theorem arguments, and epistemic utility arguments. As I mention above, these arguments are intended primarily to justify the probabilist norm, but they also serve to further characterize degrees of belief and what it means for them to be coherent. In the remainder of the chapter, I will present these arguments and explain how they help fill in some of probabilism's details. I will also briefly present two other arguments, the normative triangulation and proof in the pudding arguments, which probabilists sometimes invoke to justify probabilism in light of challenges to the standard arguments.

### 2.4.1 Dutch Book Arguments

Dutch book arguments are the most well-known—and widely criticized—justifications for probabilism. Ramsey (1964) is often credited with the first Dutch book argument, which he gives as an afterthought to his representation theorem argument for probabilism, but De Finetti (1964) gives the first full exposition of the argument.[68] The basic idea of a Dutch book argument is that, if one's degrees of belief are probabilistically incoherent, then one is vulnerable to being convinced to buy or sell a combination of bets, collectively called a "Dutch book," that will guarantee one a sure loss. Since susceptibility to a combination of sure-loss bets is a bad thing, proponents of Dutch book arguments conclude that one's degrees of belief should be probabilistically coherent.

The basic form of a Dutch book argument is as follows.[69] First we assume a connection between one's degrees of belief and one's betting behavior. Specifically, we assume that one's degrees of belief are connected to the betting quotients one deems

---

[68] See Vineberg (2011), Easwaran (2011), and especially Hájek (2009) for discussion of the origin of the Dutch book argument.

[69] My exposition of the argument and common objections to it in this section is based on Hájek (2008, 2009) and Vineberg (2011).

fair. One's betting quotient for a particular bet is the ratio of the amount the bettor would lose if their bet were incorrect over the amount of money at stake, which is the sum of the amount the bettor may lose and the amount the bettor may win. So, for example, if I were offered a bet for a \$1 stake (payout) that some claim, $A$, is true, and if I were willing to pay exactly 60 cents for the bet, my betting quotient would be $.6 = \frac{\$0.60}{\$1.00}$.[70] The nature of the connection between one's degrees of belief and one's betting quotients is not always stated precisely, so this connection is open to interpretation. At very least, one's betting quotients are supposed to express or evince—if not constitute—one's degrees of belief.

After we have assumed the connection between degrees of belief and betting quotients, we prove the Dutch book theorem, according to which, if an agent's set of betting quotients is probabilistically incoherent, then there is a set of bets the agent should be willing to take with these betting quotients that will guarantee the agent a loss, regardless of the outcomes of the events on which the wagers are made. The proof of this theorem proceeds by cases. For each axiom of the probability calculus, we show how a particular set of betting quotients (for an agent) that violates the axiom can be employed to generate a combination of bets that guarantees a sure loss (to that agent).

Next we prove the converse Dutch book theorem, according to which, if an agent's set of betting quotients is probabilistically coherent, then at least one outcome for the

---

[70] This bet counts as fair for me because, given my .6 degree of belief in A, the bet's expected payout to me = (my degree of belief that $A$ is true) × (my net gain if $A$ is true) − (my degree of belief that $A$ is false) × (my net gain if $A$ is false) = $.6 \times (\$1 - \$.6) - (1 - .6) \times \$.6 = 0$.

events on which the wager is made will not generate a loss for the agent.[71] Without the converse Dutch book theorem, there's no guarantee that probabilistic coherence will protect the agent from being Dutch booked. For, the Dutch book theorem itself does not say that probabilistically coherent sets of bets are invulnerable to Dutch book (Hájek 2009, p. 177). The converse theorem is required to establish that.

Once we've established the Dutch book theorem, we can infer that if one's degrees of belief, which are reflected in one's betting quotients, are probabilistically incoherent, then one is vulnerable to a Dutch book. But in light of the converse Dutch book theorem, we know that if one's degrees of belief are probabilistically coherent, then one is not vulnerable to Dutch book. And, since vulnerability to engaging in a combination of sure-loss bets is a bad thing, we conclude that one's degrees of belief ought to be probabilistically coherent.

The above argument affords interpretation in at least a couple of important ways. As my restatement of the argument suggests, the nature of the connection assumed between degrees of belief and betting quotients may be specified in different ways. For instance, we might contend that degrees of belief just are our betting quotients. This interpretation of the connection between degrees of belief and bets is unpalatable for several reasons. Alternatively, we might hold that one's betting quotients reflect, express, or give evidence of one's degrees of belief, though they are not, strictly

[71] The converse Dutch book theorem was proved independently by Kemeny (1955) and Lehman (1955). See Hájek (2008, 2009) for discussion. Hájek (2009, p. 177) says that the importance of the converse Dutch book theorem is often neglected in presentations of the Dutch book argument.

speaking, identical. Or we might even suggest that, while our betting quotients often reflect our degrees of belief they do not necessarily do so reliably.

Relatedly, the force of the injunction against probabilistically incoherent degrees of belief depends on whether we interpret the Dutch book arguer as holding that we should be probabilistically coherent in order to avoid actually being subject to Dutch book, or whether we should do so to avoid vulnerability to Dutch book. The significance of Dutch book arguments as justifications for probabilism depends on how one resolves these issues of interpretation.

As I note above, Dutch book arguments are not merely justifications for probabilism. They also serve as partial characterizations of degrees of belief. The arguments assume that, whatever degrees of belief are they are non-categorical investments of confidence in claims, and they afford representation by real-valued functions. But, *prima facie*, the argument does not seem to strongly constrain debates about whether we ought to be representationalists, dispositionalists, or interpretationalists about degree of belief.

Certainly, the argument has affinities with dispositionalism. On a strict interpretation of the connection between degree of belief and betting quotients, the argument relies on a form a behaviorism: one's degrees of belief are nothing other than one's betting quotients. But on more liberal interpretations, according to which betting quotients merely reflect or give evidence one's degrees of belief, Dutch book arguments seem to be compatible with interpretationalist and even representationalist accounts of degrees of belief as well. On an interpretationalist account, we would say that one's betting quotients, possibly along with other observable behavior, allow us to attribute degrees of belief to agents. And on a representationalist interpretation, we would say that our degrees of belief play a causal role in the production of our betting quotients, though they are distinct. Incoherent betting quotients are evidence that one's degrees of belief are incoherent.

While Dutch book arguments contribute to the characterization of the first tenet of probabilism, they seem only to justify the second tenet. One's degrees of belief ought to be representable as probabilities, and so ought to obey the epistemic norms the probability axioms represent; they must do so in order to avoid vulnerability to Dutch book. Dutch book arguments don't seem to provide any direct justification for the claim that beliefs come in degrees.

The nature of the justification that Dutch book arguments provide, and the notion of degree of belief coherence they characterize, depend on how we interpret the argument. If we adopt a "strict version" of the argument according to which degrees of belief are nothing other than betting quotients, and according to which we should maintain coherent betting quotients to avoid actually buying or selling a combination of sure-loss bets, then the associated characterization of probabilism is narrower, and the justification for probabilism provided by the argument is stronger, but the upshot is less interesting. If we are supposed to maintain probabilistic coherence merely in order to avoid entering into actual sure-loss bets, we would plausibly respond that we'd rather just avoid making bets (Hájek 2008, p. 799). That's easier than trying to maintain probabilistic coherence!

But if we adopt a more liberal version of the argument, a version where degrees of belief are related to, but not identical with betting quotients, we get a more interesting characterization of probabilism, but a weaker justification for the probabilist norm. For, the looser the connection between degrees of belief and betting quotients, the weaker justification the Dutch book argument provides (Hájek 2009, pp. 178-179). On this interpretation, the significance of Dutch books is that they indicate how probabilistically incoherent degrees of belief make us *vulnerable* to “pragmatic self-defeat” (Talbott 2011). If our degrees of belief are incoherent, they may lead us to behave in ways that result in worse consequences than we might have experienced if our degrees of belief were coherent. On this kind of account, Dutch book arguments serve as a kind of dramatization of the sorts of ill consequences one can incur if one’s degrees of belief are incoherent.

On either interpretation of the Dutch book argument, the kind of justification that these arguments provide seems to be pragmatic. One ought to be probabilistically coherent on pain of eventuating ill consequences due to one’s behavior. This can be seen as a positive feature of the argument insofar as it means that the Dutch book argument justification of probabilism is independent of other epistemic considerations. But it is a negative feature in that it fails to give an epistemic justification for probabilism. A way around this problem is suggested by Ramsey, who views the Dutch book argument as a dramatic device to illustrate a logical flaw in the degrees of belief of incoherent agents.[72] The idea is that if one is susceptible to a Dutch book, then one is committed to deductively inconsistent beliefs: one views the same set of bets as both fair (in the sense that they are appropriate to one's betting quotient) and unfair (in that they guarantee one a sure loss). On this view, the Dutch book argument is merely meant as a way to illustrate the results of this inconsistency. Thus, Ramsey seems to justify probabilism on the basis of accepting an epistemic norm of deductive consistency for categorical beliefs. While this view establishes an epistemic justification for probabilism, it seems to have problems of its own.

### 2.4.2 Representation Theorem Arguments

Representation theorem arguments for probabilism exemplify a general form similar to Dutch book arguments. If one's degrees of belief are not representable by a probability function, then one risks an ill consequence. In this case, the idea is that if one's degrees of belief (together with one's preferences) don't afford representation by a probability function (together with a utility function), then one's preferences fail to satisfy conditions of rationality. So, in order to avoid irrationality, we ought to have probabilistically coherent degrees of belief. The first representation theorem argument for probabilism is due to Ramsey (1964). Other prominent proponents of representation theorem arguments include Savage (1972), Jeffrey (1965), Maher (1993), and Joyce (1999).[73]

---

[72]For other "depragmatized" interpretations and versions of the argument, see, for example, Christensen (1996), Christensen (2001) and Howson & Urbach (2006).

[73]It's worth pointing out that Joyce's (1999) treatment of representation arguments overcomes some of the usual problems discussed below. Joyce's account takes the

The basic form of representation theorem arguments for probabilism is roughly as follows.[74] First we assume (or argue) that a version of the standard axiomatization of rational preference is correct. That is, the preference axioms provide an appropriate representation of the preferences over options or outcomes (or goods) for rational agents, and, thereby, they represent norms for rational preference. Different versions of the axioms offer slightly different norms, but they tend to include requirements that preferences should be transitive, reflexive (and perhaps complete), and impartial (with respect to the addition of independent options or outcomes).

Next we appeal to a representation theorem, according to which, one is representable as satisfying the axioms of rational preference if and only if one can be represented as an expected utility maximizer by means of a utility function and a probability function. Such representation theorems are acknowledged to be beyond doubt given the assumption of appropriate axioms for rational preference.[75]

---

agent to have both preferences and comparative confidence levels, and he presents his preference axioms in terms of these two relations. Then, his representation theorem shows that his confidence-preference axioms are satisfied by an agent just in case the comparative confidence relation is representable by a probability function, the preference relation is representable by a utility function, and the preference orderings that satisfy the confidence-preference axioms correspond to the orderings among the resulting expected utilities. Thanks to Jim Hawthorne for pointing this out.

[74] My presentation of the argument and its interpretation is based on the expositions in Zynda (2000), Easwaran (2011), and especially Hájek (2008).

[75] In Chapter 4, I will talk about a different kind of representation theorem, according to which if an agent's degrees of belief are representable with a comparative confidence relation satisfying certain axioms, then the agent's degrees of belief can also be represented with a probabilistic credence function (or set thereof).

Given the representation theorem, we then infer that if one cannot be represented as an expected utility maximizer *via* a utility function and a probability function, then one cannot be represented as satisfying the axioms of rational preference. So, if one can't be represented with a probability function (usually interpreted as a credence function), then one can't be represented by the rational preference axioms. That is, one is irrational. So, to avoid irrationality, we ought to maintain probabilistically coherent degrees of belief.

There are a three of important points of this argument that merit attention and interpretation. First, the argument seems to operate by sleight of hand (Hájek 2008, pp. 803-804). We draw a conclusion about norms for degrees of belief from assumptions about the formal representation of preferences. We don't mention degrees of belief at all until the conclusion of the argument. To draw this conclusion, we tacitly assume that the probability function that is required by the representation theorem is a *credence* function. We assume that being representable by a probability function means that we have degrees of belief, which are representable with a probability function. Another way to put it is that the argument is tacitly realistic about the probabilistic representation. Without this assumption we cannot draw any conclusions about probabilism from assumptions about rational preferences.[76]

---

[76] Indeed, Zynda (2000) shows that, if an agent's preferences satisfy the preference axioms, then they can be represented as a non-expected utility maximizer via a non-probability function and a non-utility function. This suggests that there's no reason to favor the probabilistic representation over the non-probabilistic representation, so the representation theorem argument fails to support its conclusion, the probabilist norm (Hájek 2008, pp. 805-806).

Second, it’s important to note that the argument assumes the correctness of a version of the rational preference axioms. If the axioms don’t provide an appropriate formal representation of the preferences of rational agents and if they don’t provide appropriate norms for preference, then the argument fails. It’s also important to note that this assumption is akin to the assumption that representation by a probability function implies that one has coherent degrees of belief. The assumption of the appropriateness of the rational preference axioms is an assumption that representation according to the axioms implies that one has rational preferences. Are the preference axioms any better justified than the probabilist norm?

Like Dutch book arguments, representation theorem arguments provide a further characterization of the notion of degree of belief. The most obvious way that representation theorem arguments provide a further characterization of degree of belief is by drawing connections between degree of belief and preference: if one’s preferences are rational, then one’s degrees of belief must also be rational.

In light of this connection, it seems natural to adopt a dispositionalist—even behaviorist—interpretation of degree of belief. On this kind of interpretation, one’s degrees of belief are entirely explained in terms of one’s behavior. The representation theorem places additional constraints on how the behavior that constitutes one’s degrees of belief must relate to the behavior that reveals one’s preferences.

The representation theorem argument also fits well with an interpretationalist account of degree of belief. On this sort of account, degree of belief is understood as a theoretical entity appealed to explain behavior in conjunction with other theoretical entities, like preference. On an interpretationalist view, the representation theorem argument further specifies the way in which we must interpret behavior in terms of degree of belief and preference.

While the representation theorem argument accords well with dispositionalism and interpretationalism, it can also be understood in accordance with a representationalist account of degree of belief. On this kind of account, the representation theorem argument helps specify the causal relationship between, namely, degree of belief and preference.

Like Dutch book arguments, representation theorem arguments seem to provide justification only for the second tenet of probabilism. It seems only oriented at justifying treating the probability axioms as representing norms for degree of belief. It does not offer any direct justification for the claim that beliefs come in degrees. Though one might argue that it could give indirect justification for the first tenet on the basis that, if one has rational preferences, one can be represented with probability function. One might take this as an indication that rational preferences require that agents have doxastic states that come in degrees.

The justification that we get for probabilism in a representation theorem argument comes from the appropriateness of the axioms for rational preference and the assumption that the probabilistic representation they require is representative of states the agent actually has. So, the justification for probabilism that we get from the representation theorem argument is entirely parasitic on the justification for the preference axioms. So, in order to fully understand the justification that RTAs provide for probabilism, we must also investigate the justification on the axioms for rational preference. That's outside the scope of the current project.

Additionally, the justification for probabilism depends on the assumption that probabilistic representation of the agent required by the representation theorem is true to states the agent is actually in. But there may be reason to doubt this assumption. It may be that agents are in states similar to, but clearly distinct from degrees of belief that afford probabilistic representation. If this is the case, then the representation theorem argument doesn't justify probabilism. It merely justifies the value of being probabilistically representable.

Supposing that these assumptions are appropriate, and the representation theorem argument for probabilism goes through as intended, it is important to note that the justification it purports to provide for probabilism is non-epistemic. For according to the representation theorem argument, probabilism is justified by some features of rational preferences. As in the case of the Dutch book argument, the non-epistemic nature of the justification for probabilism could be construed as a positive feature of the representation theorem argument. It provides a non-epistemic ground for epistemic norms. But it can also be construed as a negative feature. It seems like we should be able to come up with an epistemic justification for our epistemic norms.

### 2.4.3 Epistemic Utility Arguments

As I explain above, Dutch book arguments and representation theorem arguments for probabilism are usually seen as providing merely pragmatic, non-epistemic justification for probabilism. Many probabilists want to justify probabilism on what they see as purely epistemic grounds. Now I will briefly present two kinds of arguments, often called "epistemic utility arguments," that are intended to provide this kind of justification. These arguments attempt to identify an epistemic value, like accuracy, and then show that our degrees of belief should be probabilistically coherent in order to better satisfy this value.

#### 2.4.3.1 Calibration Arguments

Calibration is a measure of how well one's degrees of belief match corresponding relative frequencies. One's credence function is perfectly calibrated, if for each $r$ between 0 and 1, the proportion of true claims to which one assigns a degree of belief $r$ has the value $r$. So, for example, one's degrees of belief regarding the chance of rain are perfectly calibrated if it rains on $10\%$ of the days for which one's degree of belief that it will rain is .1, it rains on 75% of the days for which one's degree of belief that it will rain is .75, and so on.[77] The first calibration arguments for probabilism are due to van Fraassen (1983) and Shimony (1988).[78]

The basic idea of calibration arguments for probabilism is to prove the following kind of claim: if one's degree of belief function is probabilistically incoherent, then there is some probabilistically coherent function that is better calibrated. Since, calibration is a good thing, the argument goes, our degrees of belief ought to obey the probability axioms.

---

[77]This example is borrowed from Hájek (2008, p. 807).

[78] My presentation of the argument is based on the presentations in Joyce (2004) and Hájek (2008).

The calibration argument begins with a proof of the calibration theorem, according to which, if one's degree of belief function violates the probability axioms, then there is some probability function that is better calibrated than one's degree of belief function under every deductively consistent assignment of truth-values to claims (Hájek, 2008, p. 807). Then, on the assumption that perfect calibration is epistemically valuable, we infer that a degree of belief function that is better calibrated is superior to less well-calibrated degree of belief functions. Thus, we conclude that, since degree of belief functions that are probabilistically coherent are better calibrated than incoherent degree of belief functions, our degrees of belief ought to be probabilistically coherent.

The characterization of degree of belief that we get from calibration arguments is most naturally construed representationally. On this view, an agent's degrees of belief in claims are mental representations that ought to be attuned to the frequencies with which the claims in question are true. While calibration arguments fit well with representationalist interpretations of degree of belief, they also seem to be compatible with dispositionalist and interpretationalist construals of degree of belief. On a dispositionalist interpretation, calibration arguments show how the dispositions that constitute degree of belief ought to be attuned to the relative frequencies of the truth of the claim. For example, on a behaviorist interpretation, one's actions on the assumption that $A$ ought to accord well with the relative frequencies of the truth of $A$. On an interpretationalist account, calibration arguments characterize how the degrees of belief attributed to rational agents ought to represent the agents' observable behavior as conforming to the relative frequencies with which the claims in which they invest degree of belief are true.

Like the other arguments for probabilism, calibration arguments are aimed primarily at justifying the second tenet of probabilism. They are intended to provide non-pragmatic, epistemic justification for probabilism by showing that coherent degrees of belief are better calibrated to the world in a certain sense. The justification that calibration arguments provide depends crucially on the assumption that perfect calibration is an epistemic good in its own right. However, several authors have challenged the assumption that calibration is a genuine epistemic good.[79]

**2.4.3.2 Accuracy Arguments**

The basic idea underlying accuracy arguments for probabilism is to establish (via a mathematical proof) that probabilistically coherent credence functions are more accurate than incoherent functions (according to some measure of accuracy). Since, we favor accurate degrees of belief, our degrees of belief ought to be probabilistically coherent. The first accuracy argument for probabilism is due to Joyce (1998). In recent years, accuracy arguments for probabilism and other tenets of Bayesianism have taken off.[80]

Accuracy arguments for probabilism begin with a proof of the gradational accuracy theorem, according to which, if one’s degree of belief function violates the probability axioms, then there is a probability function that is strictly more accurate under every

[79]See, for example, Seidenfeld (1985), Joyce (1998), and Hájek (MSa).

[80] See Pettigrew (2011) for a survey.

logically consistent assignment of truth values to claims.[81] That is to say, if one's degree of belief function is probabilistically inconsistent, there is another probabilistically consistent degree of belief function that is more accurate, for every possible way the world could be. That is, any incoherent degree of belief function is accuracy-dominated by a coherent function. Since being accuracy-dominated is a bad thing, we conclude that our degrees of belief ought to be probabilistically coherent.

Like calibration arguments, accuracy arguments for probabilism fit naturally with representationalist accounts of degrees of belief. On this kind of view, one's degrees of belief are mental representations that ought to accurately reflect the way the world actually is. In order to avoid inaccuracy, one ought to maintain probabilistically consistent degrees of belief.

Accuracy arguments also seem to be compatible with dispositionalist and interpretationalist accounts of degree of belief. On a dispositionalist account, accuracy arguments require the dispositions that constitute degrees of belief to accurately reflect the way the world is. If the dispositions in question are behavioral, then one ought to behave as if the true claims are true. If the dispositions are mental, then they ought to represent the world accurately, or in accord with other accurate representations. On an interpretationalist account, accuracy arguments are construed as the injunction to behave so that one can be rationally attributed accurate degrees of belief.

---

[81] My presentation of the gradational accuracy argument is based on Hájek (2008) and Pettigrew (2011).

Like the other arguments for probabilism I have considered, accuracy arguments are aimed only at establishing the second tenet of probabilism. And, like calibration arguments, they purport to provide non-pragmatic, epistemic justification for probabilism.

**2.4.4 Normative Triangulation and Proof in the Pudding Arguments**

In addition to the three standard kinds of arguments for probabilism just surveyed, probabilists sometimes also offer two additional kinds of arguments for probabilism, which I call the "normative triangulation"[82] and "proof in the pudding" arguments. Unlike the standard arguments, these arguments are not intended to characterize degrees of belief, or what it means for them to be coherent. Indeed, in some cases,[83] these arguments are meant to compatible with primitivism about *degree of belief*.

According to the normative triangulation argument, while each of the standard arguments for probabilism has important flaws, they each also help characterize degrees of belief and what it means for them to be coherent, and they help tie probabilism to our other commitments about degrees of belief.[84] Degrees of belief *are* connected to betting behavior, and they *do* bear a normative connection to preferences, and (at least sometimes) it *is* good for them to be calibrated and accurate. As Hájek (2008, p. 816) puts it, together the standard arguments help us "triangulate" to

---

[82] Following Hájek (2008, p. 816).

[83] Like Hájek (2008).

[84] See Christensen (2001), Eriksson & Hájek (2007), and Hájek (2008) for this kind of view.

probabilism. This quote from Christensen (2001, p. 375),[85] writing of Dutch book and representation theorem arguments, sums up the view nicely:

> Neither one comes close to being a knock-down argument for Probabilism, and non-probabilists will find contestable assumptions in both. But each one, I think, provides Probabilism with interesting and non-question-begging intuitive support. And that may be the best one can hope for, in thinking about our most basic principles of rationality.

According to the proof in the pudding argument, probabilism is also justified by its theoretical fruits. The idea is that probabilism provides the foundation of a very rich research program, namely subjective Bayesianism, which shows great promise for providing a unified theory of doxastic rationality, practical rationality, and scientific confirmation. It also helps resolve several troubling philosophical problems in these areas. And it provides an excellent characterization of a highly useful psychological and epistemological notion, namely degree of belief.

### 2.5 Conclusion

In this chapter, I presented a typical statement of probabilism and the most prominent (and widely discussed) options for filling in the details. I presented probabilist characterizations of degrees of belief, and justifications for positing degrees of belief. I also presented the probability calculus, its interpretation as a source of degree of belief norms, and the standard arguments probabilists use to justify the probabilist norm. This presentation was intended primarily to acquaint the reader with probabilism. However,

[85] Quoted in Easwaran (2011, p. 318).

this presentation as also intended to demonstrate the need for probabilists to fill in the details of their view.

In Chapter 3, I will argue more explicitly that probabilists need to fill out the foundations of their view. I will show how three common objections to probabilism arise from probabilism's typical under-articulation. I will argue that probabilists must spell out the methodological and normative foundations of their view in a unified, systematic way in order to avoid the most pressing objections. Then, in Chapters 3 and 4 I will develop an account of these foundations that is more detailed, and more satisfactory than those available to date.

## 3 Probabilism & Its Discontents

In Chapter 2, I presented the typical two-tenet characterization of probabilism along with many of the main options for filling it out. This presentation was intended to characterize probabilism, the subject of this dissertation, but it was also intended to demonstrate some of the ways that probabilism's details need to be filled in. Thus, Chapter 2 was intended to familiarize the reader with probabilism and plant the seed of an argument for the claim that probabilists need to fill out their view and articulate its foundations.

However, the mere fact that probabilism's typical characterization leaves many questions unanswered doesn't, on its own, give us very compelling reasons to attempt to develop an account of its foundations. Probabilists may have good reasons to focus on applications of their view and keep the foundations implicit. For probabilism is a pretty simple, intuitive view, and it is easy to get it up and running as it applies to particular problems with only a minimal understanding of the philosophical foundations of the view. Probabilism is also a very flexible view. It has applications to everyday epistemological problems as well as problems in philosophy of science and decision theory. Probabilists may be wary of compromising the intuitiveness, simplicity, and flexibility of the view by weighing it down with foundational baggage. Hence the "probabilist formula" in the probabilist literature that I mentioned in the introduction: give the two-tenet sketch of probabilism and start churning out the formulas for specific applications.

In this chapter, I will provide additional reasons to think that probabilists need to articulate unified foundations for their view. My approach will be to draw upon crucial components of the dialectic between probabilists and their critics. I will present three main sources of objections to probabilism: the probabilist scheme for formally representing degrees of belief, probabilism's apparent requirement that an agent be strongly "opinionated," and probabilism's apparent requirement that an agent be "logically omniscient." Critics of probabilism—often probabilists themselves of one stripe or another—object to these aspects of probabilism for three main reasons: (1) they are psychologically implausible, (2) they are excessively demanding or idealistic, and (3) they are of doubtful value. Over the years, probabilists have developed various responses to these criticisms—for example, probabilism relies upon both methodological and normative idealizations; probabilists should be local modelers (and local models are not so overwhelmingly idealistic), etc. This dialectic raises several deep questions about probabilism's normative and methodological foundations. Furthermore, while probabilists' responses to these objections may begin to reveal answers to these foundational questions, they have only done so in a piecemeal way, leaving it unclear whether any single unifying coherent account of the foundations can be supplied. So, I contend that probabilists need to provide unified foundations for their view. Only by doing so can they provide a clear account of the content and aims of the view, showing that a coherent rebuttal of objections can stand up to scrutiny. Furthermore, such a unified account can contribute to reining in illicit applications of the view based on misunderstandings of its content and aims.

In §3.1 I will present details of the dialectic between probabilists and their critics. In §3.2 I will draw out the main critical themes and responses from this dialectic. I will also identify several deeper problems suggested by the usual objections, and I will classify strategies probabilists may employ to respond to them. In §3.3, I will explain how some of the usual probabilist responses are at cross-purposes, and I will argue that probabilists must provide unified foundations in order to avoid such internal conflicts in the articulation of their view.

### 3.1 The Dialectic: Probabilists & Their Critics

Probabilism—and Bayesianism, generally—have been subject to several important objections since they first started gaining popularity. In this section, I will present three of the main sources of objections to probabilism. In doing so, my aim is not to suggest that other objections to probabilism are unimportant, or that probabilist responses to those objections are entirely satisfactory. I have chosen to focus on the following objections and responses because of the centrality of the foundational issues the attendant dialectic reveals.

First, I will present the dialectic concerning probabilism's scheme for representing degrees of belief formally. Then I will consider objections and replies concerning probabilism's apparent requirements that agents be opinionated and logically omniscient.

#### 3.1.1 The Formal Representation Scheme

Recall the probabilist scheme for formally representing degrees of belief presented in Chapter 2. Probabilists represent the claims in which an agent invests degrees of belief with the sentences of a logical language, $\mathcal{L}$. The atomic sentences of $\mathcal{L}$ represent the atomic claims in which the agent invests confidence, and $\mathcal{L}$ is closed under the usual logical operations, so it contains all of the truth-functional combinations of these atomic sentences. The agent's degrees of belief are represented with a pair of functions defined on $\mathcal{L}$, an unconditional credence function, $cr(\cdot): \mathcal{L} \to [0,1]$, and a conditional credence function, $cr(\cdot \mid \cdot): \mathcal{L} \times \mathcal{L} \to [0,1]$. The outputs of these functions represent the strength of the agent's degree of belief in the claims represented by the input sentences of $\mathcal{L}$. Thus, each of the agent's degrees of belief is represented with a precise numerical value. Values closer to 1 represent stronger degrees of belief; values closer to 0 represent weaker degrees of belief (that is, stronger degrees of belief that the claim is false). The probabilist norm requires the credence functions that represent an agent's degrees of belief to be probability functions.

I will present three objections to this formal representation scheme. The first—perhaps naïve—objection holds that probabilism is implausible because it requires agents to have numbers—numerical degrees of belief—"in their heads." The second, more sophisticated objection holds that agents' doxastic attitudes are not even *representable* with precise numerical credence functions. The third objection holds that since humans are in general not very good intuitive judges of probability, any *probabilistic* representation of degrees of belief is implausible.

#### 3.1.1.1 Numbers in the Head?

The *no numbers in the head* objection holds that probabilism's formal representation scheme seems to assume that we have uniformly explicit, precise numerical degrees of belief "in our heads," but that this assumption is simply false. We rarely—if ever—have precise, explicitly numerical degrees of belief, and it doesn't seem like we could or should have them all the time. Our phenomenology does not reveal many conscious numerical degrees of belief. We sometimes attach percentage estimates or other numerical quantities to our degrees of belief. But these percentages are just *estimates* of the strength of our degrees of belief; they aren't inherent to the degrees of belief themselves. Sometimes it is difficult to precisely quantify the strength of our confidence. In some cases, such difficulties are instances of the general problem of fixing our doxastic attitudes in relation to our evidence and our other doxastic attitudes. In other cases, our degrees of belief defy precise numerical estimation because they are vague or imprecise.

It doesn't seem like we have subconscious numerical degrees of belief either. For one thing, it's not clear what this would mean. We'd have to appeal to a theory of what degrees of belief are *qua* mental state in order to make sense of this. Additionally, it would be strange for us to have subconscious numerical degrees of belief given how rarely our conscious degrees of belief are numerical, and how difficult it is sometimes to quantify our confidence. Given the difficulties we sometimes encounter in quantifying our degrees of belief, it seems like we couldn't have uniformly precise numerical degrees of belief. And it's also not clear why we would want to.

The *no numbers in the head* objection is not one that you'll often encounter in the literature.[86] It depends on a conflation of degrees of belief with their formal representations. Or, more charitably, it requires a strong realism about probabilism's formal representation scheme. As such, it is a bit naïve and a bit of a straw man. However, it is the sort of objection one often hears from those encountering probabilism for the first time. And it is understandable given that probabilists rarely characterize their view in enough detail to preclude this interpretation. Indeed, the way many probabilists talk and write about degrees of belief and their numerical representations invites the conflation of the two. So, while the *no numbers in the head* objection may be naïve, it raises an important question, namely, what exactly does it mean for our degrees of belief to be represented with real-valued credence functions*?*

When probabilists encounter the *no numbers in the head* objection, they typically explain the problem away without much fanfare. They explain that probabilism's formal representation scheme doesn't require agents to have numerical degrees of belief *in the head*. The numerical representation of degrees of belief is just that—a *representation*. This doesn't mean that agents have explicitly numerical degrees of belief—whether consciously or subconsciously. To say that an agent $\alpha$'s degree of belief in a claim $A$ can be represented by a credence function with an output .999 ($cr_{\alpha}(A) = .999$) is not necessarily to say that $\alpha$ has a mental state that attaches .999 to $A$. Minimally, it means

[86] Harman (1986, p. 22) considers a version of probabilism that posits explicitly numerical degrees of belief in agents, and he raises this objection, but he acknowledges the implausibility of such a view and does not attribute it to probabilists. See also Zynda (2000, p. 50) for a brief discussion.

that $\alpha$ is nearly certain that $A$ is true. For the representation scheme holds that numbers closer to 1 represent stronger degrees of belief, and .999 is pretty close to 1.[87] Depending on the specific account of the numerical representation, the precise credence value .999 may have additional significance. For example, it may indicate that $\alpha$ would deem it fair to engage in a bet (on either side) that pays out \$999 (to one who bets against A) if $A$ is false for a gain of \$1 (to one who bets for $A$) if A is true. But it does not guarantee that $\alpha$ has an explicitly numerical degree of belief in $A$. Numerical credence functions are meant to capture the gradational structure of our degrees of belief, but they may not do so perfectly. The precise numerical representation of degrees of belief is mostly a technical convenience for ease of representation and computation.

These responses provide a superficial resolution to the *no numbers in the head* objection: probabilism doesn't make the objectionable assumption that agents have explicitly numerical degrees of belief; it merely *represents* agents' degrees of belief numerically. However, these responses do not resolve the deeper issue of exactly what it means for one's degrees of belief to be representable with numerical credence functions or for one's credence functions to be probability functions. To resolve these deeper issues probabilists must provide a more detailed account of what it means to have degrees of belief and what the gradational structure of degrees of belief is really

[87] The language of "closeness" suggests that the numbers in the probabilist representation scheme have more than mere ordinal significance. For there are uncountably many real numbers between .999 and 1. So, in an ordinal sense, .999 is still infinitely far from 1.

like. Only then can we fully understand what it means for our degrees of belief to be represented with sharp, real-valued credence functions. Only then can we understand to what extent this representation is faithful and to what extent it is distorted for the sake of simplicity and computational tractability.

The *no numbers in the head* objection also raises issues about whether, and how much, the abilities and limitations of actual human agents should constrain probabilist theorizing, including the probabilist formal representation scheme. If probabilism is meant to guide or provide a standard of evaluation for human cognition, perhaps its requirements should be constrained by the capabilities of actual humans. If humans couldn't have degrees of belief as richly structured as the real line in the unit interval, perhaps probabilism's formal representation scheme is inappropriate. Probabilists may be able to avoid this objection without grappling with the reasons that support it by insisting that probabilism is a normative theory, and as such is not beholden to the limitations of humans. But this raises further questions about the nature and sources of the normativity of probabilism's requirements.

#### 3.1.1.2 Numerical ***Representability***?

So, explaining that probabilism doesn't assume that we have sharp numerical degrees of belief *in the head* doesn't get probabilists off the hook. There's still a lot to explain about what it means for our degrees of belief to be representable with sharp numerical credence functions. But even if we grant the probabilist response to the *no numbers in the head* objection and we set aside most of these questions about the nature of degrees of belief, their gradational structure, and how faithfully credence functions represent it, there are other reasons to doubt the appropriateness of probabilism's formal representation scheme. Critics offer two more objections to the probabilist representation scheme, which I will call the "psychological objection" and the "evidential objection," respectively. According to the psychological objection, the precise numerical representation of degrees of belief is psychologically implausible: we don't seem to have precise numerical degrees of belief; and it doesn't seem like we could. According to the evidential objection, sometimes our evidence requires us to adopt imprecise degrees of belief that cannot be faithfully represented with sharp numerical credence functions. These objections are refinements of the *no numbers in the head* objection, and they show that concerns about psychological plausibility and normative appropriateness persist even on a more sophisticated interpretation of the probabilist representation scheme.

**3.1.1.2.1 The Psychological Objection**

Critics propose the psychological objection for reasons similar to those offered in support of the *no numbers in the head* objection. There are many claims for which we do not seem to have sharp degrees of belief that admit of precise numerical representation. Furthermore, it doesn't seem like we could have uniformly sharp degrees of belief. So, it doesn't seem appropriate to require agents to have degrees of belief that admit of numerically precise representation.[88]

---

[88] For discussion, see, for example, Kaplan (2002, p. 435), Joyce (2011, pp. 282-283), and Christensen (2004, pp. 144-150).

Probabilists offer three responses to the psychological objection. First, they explain that our lack of conscious awareness of the sharpness of our degrees of belief, and our inability to articulate our degrees of belief with numerical precision do not mean that our degrees of belief are imprecise, and that, therefore, they may not be modeled with numerical precision.[89] Our degrees of belief simply might not be available to introspection. They might be subconscious mental representations or they might be behavioral dispositions of which we have limited awareness. That is, our degrees of belief might be precise despite our inability to recognize them as such.

Critics respond in turn that our lack of awareness of the sharpness of our degrees of belief and the difficulty of assigning precise numerical estimates of the strength of our confidence should cast doubt upon characterizations of degrees of belief as subconscious mental representations or behavioral dispositions that might disagree with our conscious feelings of confidence.[90] This idea is that if an account of degrees of belief (as subconscious mental representations or behavioral dispositions) attributes degrees of belief to agents that conflict with their conscious feelings of confidence, this is a reason to doubt that account of degrees of belief.

Second, probabilists note that their numerical representation of degrees of belief is a normative idealization.[91] Probabilism is an account of ideal rationality. It's a theory of

[89] For an example of such a response, see Eells (1982, p. 41 ff.) cited in Kaplan (2002, p. 435, n. 2) and Christensen (2004, p. 144).

[90] Kaplan (2002, p. 435) attributes this view to Goldman (1986, pp. 326-328).

[91] For example, see Christensen (2004, pp. 146-147).

how agents *ought* to apportion their degrees of belief, or, alternately, a theory of how it is *good* for agents to apportion their degrees of belief. As such, probabilism is not constrained by the abilities of actual epistemic agents. The facts that actual agents rarely have sharp degrees of belief that afford faithful numerical representation, and that agents like us do not seem capable of having uniformly precise degrees of belief do not cast doubt on the claim that agents ought to have precise degrees of belief or that it would be good for agents to have precise degrees of belief.

Third, probabilists also explain that in addition to being a normative idealization, the numerical representation of degrees of belief is also an instance of so-called "Galilean idealization"—the practice of intentionally misrepresenting the modeled phenomenon in order to achieve a more tractable model.[92] Representing degrees of belief with precise numerical values between 0 and 1 allows probabilists to employ the elegant, fecund, and decidable[93] probability calculus rather than a messier, more complicated, but more realistic formal modeling framework. The representation of an agent's degrees of belief by means of a unique pair of probabilistic credence functions makes it easy to compute the coherence of the agent's degrees of belief. And this idealization is close enough for most real purposes. Less precise numerical representations and qualitative formal representations would be less computationally convenient and would gain little benefit for practical purposes. Together, these

---

[92]See Christensen (2004, pp. 145-146) for an example of a probabilist who gives this sort of account of the idealization implicit in probabilistic modeling degrees of belief. For a more detailed account of Galilean idealization, see Weisberg (2007, p. 640).

[93] Assuming that the underlying logic is decidable. See Fitelson (2008).

probabilist responses provide plausible defense of the sharp probabilistic representation of degrees of belief. Some probabilists will not want to endorse the first response in order to avoid disconnecting our degrees of belief from our conscious awareness. But many probabilists avail themselves the second and third responses.

These responses seem to allow probabilists to avoid the psychological objection, but, in doing so, they raise further questions. Like the *no numbers in the head* objection, the psychological objection raises the issue of how realistically we ought to interpret the numerical representation of degrees of belief, and how much the normative significance of probabilism should be constrained by the abilities of real agents. Probabilists avoid the psychological objection by allowing that the unique numerical representation of degrees of belief may be an oversimplification for the sake of computational convenience, and by denying that the abilities of real agents impose a firm constraint on normative theorizing. But this response prompts critics to raise once more the questions of how much technical simplification is acceptable and about the extent to which the abilities of real agents should provide constraints on probabilist norms.

**3.1.1.2.2 The Evidential Objection**

The response that probabilism employs various idealizations seems to save it from the psychological objection, but at the cost of attributing to probabilism a misleading nomenclature associated with its formal representation of degrees of belief, and an implausible account of epistemic normativity. The evidential objection to the precise numerical representation of degrees of belief poses a more difficult challenge. According to the evidential objection, sometimes our evidence is “incomplete, imprecise, or equivocal" (Joyce 2011, p. 283). In such cases, our evidence requires us to adopt incomplete, imprecise, or equivocal degrees of belief that cannot be faithfully represented with unique, sharp probabilistic credence functions.[94]

Joyce provides the following example of such a case (2011, p. 283):

> **Black/Grey Coins.** An urn contains a large number of coins which have been painted black on one side and grey on the other. The coins were made at a factory that can produce coins of any bias $\beta: (1 - \beta)$ where $\beta$, the objective chance of the coin coming up black, might have any value in the interval $0 < \beta < 1$. You have no information about the proportions with which coins of various biases appear in the urn. If a coin is drawn at random from the urn, how confident should you be that it will come up black when tossed?

In this case it seems that adopting any sharp credence that the coin will come up black would be to disrespect the evidence. Any point-valued credence will be more precise than the evidence supports. Thus, it seems that the sharp numerical representation of degrees of belief required by probabilism can be at odds with various evidential requirements.[95] So, it's normatively inappropriate for probabilism to require the representation of ideal epistemic agents with precise numerical degrees of belief. For, in the face of such evidence, even computationally and mnemonically ideal, but evidentially limited epistemic agents would adopt imprecise degrees of belief.

---

[94] See Christensen (2004, pp. 147-150) and Joyce (2011) for discussion.

[95]It may be important to distinguish the notion of precision or sharpness from the notion of uniqueness. Many epistemologists have argued recently that for any set of evidence, there is a unique doxastic attitude that it supports. This thesis seems to be consistent with the claim that evidence may require an agent to adopt a non-sharp degrees of belief, so long that the evidence supports a unique credence. So, perhaps the credence required by the evidence in the black/grey coins case is best modeled by a set of probability functions that yield outputs throughout the $(0,1)$ interval. This credal state is non-sharp, but nonetheless unique.

Thus, the probabilist cannot brush off the evidential objection to the numerical representation of degrees of belief by explaining the normative idealization of probabilism. It seems that even ideal agents should sometimes have imprecise degrees of belief.

This leaves the probabilist with the response that the precise numerical representation of degrees of belief is a Galilean idealization—a simplification for the sake of computational tractability. This response might save the probabilist from the evidential objection, but it does so at a cost. For it raises the issue of how much simplification is acceptable in the formal representation of degrees of belief. It seems that the precise numerical representation of degrees of belief is unacceptable because its computational convenience puts it at odds with evidential norms. This suggests that the probabilist formal representation of credence ought to be revised to allow for imprecise, incomplete, or equivocal degrees of belief. Thus, the evidential objection pits the normative aspirations of probabilism against its technical convenience.

A common probabilist response to the evidential objection is to grant that the sharp numerical representation of degrees of belief is sometimes inappropriate, and to augment the traditional probabilist formal modeling framework by allowing that an agent's doxastic state may be best modeled in a given instance by a set of credence functions which together represent the coarse-grained range of an agent's imprecise degrees of belief.[96] As we saw in Chapter 2, qualitative probabilists even suggest that

[96] See Walley (1991), Christensen (2004, pp. 147), White (2009), and Joyce (2011), among others, for this kind of response.

agents' doxastic states should be represented with qualitative comparative confidence relations, which may in turn be modeled with probabilistic credence functions or sets thereof.[97] These qualitative probabilist modeling frameworks allow us to model agents with incomplete, equivocal, and imprecise degrees of belief.

The evidential objection and the probabilist responses to it raise familiar issues concerning how to interpret probabilist formal models, how realistically the formalism should be construed, how much simplification is appropriate in a formal model, how the abilities of real agents constrain normative theorizing, and how the requirements of probabilism interact with other epistemic norms, such as evidential norms. Probabilists must answer these questions in order to fully respond to the above concerns about the sharp numerical representation of degrees of belief.

#### 3.1.1.3 Issues Raised by Poor Human Probability Judgment

Even if we grant that precise numerical representation of degrees of belief is appropriate (for a wide range of cases, perhaps), we might still worry about the *probabilistic* representation of degrees of belief. We might worry that, because humans are poor intuitive judges of probability, our degrees of belief are unlikely to be faithfully probabilistically representable, and, therefore, the probabilistic representation of degrees of belief is inappropriate.[98]

Researchers in the cognitive biases and heuristics tradition have amassed compelling evidence that the intuitive probability judgments of humans depart from the

[97] See Hawthorne (2009; MS).

[98] This seems to be the suggestion in Pollock (MS).

dictates of probability theory in significant and systematic ways.[99] For instance, they have found in experiments that humans often ignore prior probability distributions when reasoning about the prevalence of phenomena within a particular population (Tversky & Kahneman 1974). They have also found that humans often fail to conform to the probabilistic rule that the probabilities of entailed sentences must be at least as high as those that entail them (Tversky & Kahneman 1983). Since humans often make such fundamental probabilistic errors, it seems that humans' degrees of belief are unlikely to often be faithfully representable with probabilistically coherent credence functions.

Some philosophers see the significant and systematic descriptive inadequacy of the probabilistic representation of degrees of belief as a reason to doubt its appropriateness for a normative account of doxastic rationality. These philosophers worry that probabilistic representability cannot be epistemically good or obligatory, if humans systematically fail to live up to that standard.

The primary probabilist response to this worry is that probabilism is a *normative* account of doxastic rationality. It is not a description of the way agents actually invest confidence in claims. The fact that humans are poor intuitive judges of probability, and that our degrees of belief do not often admit of probabilistic representation, is compatible with the claim that probabilistic representability is an appropriate normative requirement. The above mentioned research on human reasoning also shows that

---

[99] For overviews of this literature at different stages of its development, see, for example, Kahneman et al. (1982), Gilovich, et al. (2002), and Kahneman (2011) and Pohl (2012).

humans are poor deductive reasoners by the standards of formal deductive logic. Does this make formal deductive logic an inappropriate standard for human deductive reasoning? Some probabilists including Christensen (2004, pp. 154-164) and Titelbaum (2013, pp. 60-75) explain further that probabilism is an evaluative—as opposed to deontological—account of doxastic rationality. Its requirements describe epistemic ideals that agents may or may not be able to meet, rather than epistemic obligations that agents ought to live up to. As such, it is not constrained by the abilities of actual agents.

This response allows probabilists to insist upon the appropriateness of probabilistic representations of degrees of belief. But it raises the issue of exactly what the normative significance of probabilism's requirements is supposed to be. Are these requirements deontological or evaluative? Or do they have some other kind of normative force? What is their significance for humans? How do these requirements relate to other epistemic norms?

### 3.1.2 Opinionation

As we saw in Chapter 2, probabilism seems to include an opinionation requirement. Kolmogorov's axioms specify that probability functions are total functions: they map each input to some output or another. So, a probabilistic credence function maps each sentence of the modeling language to some real number between 0 and 1. Translated into a norm for degrees of belief, this seems to be a requirement that an agent must have a unique, precise degree of belief toward every claim represented in the modeling language. Recall that this includes every logical (or set-theoretic) combination of the atomic claims represented in the modeling language.

However, that's not all there is to the opinionation requirement. For probability functions represent the strength of an agent's confidence in a univocal numerical scale. That is, the strength of an agent's confidence in each claim is represented with a real number between 0 and 1. And we can appeal to these numbers to compare differences in the agent's confidence in claims. Implicitly, this numerical representation makes every pair of claims comparable on a single scale. So, in addition to requiring that you have some attitude or other to every claim, opinionation also seems to require that you should be able to compare the strength of your confidence in every claim to every other with an arbitrarily high degree of precision. So, the opinionation requirement seems to demand that we should have a huge number of precise doxastic attitudes, all of which are comparable to each other.

Critics of probabilism express three main worries about the opinionation requirement. First, they worry that the opinionation requirement is inappropriate because it requires us to have more doxastic attitudes than we could possibly have. Second, they worry that opinionation requires us to have more doxastic attitudes than we would want to have—it requires us to clutter our minds with irrelevancies. Third, it requires us to have more degrees of belief than we could possibly hope to maintain coherently—we could not conduct the necessary computations to maintain coherence.

### 3.1.2.1 Can We Have So Many Doxastic Attitudes?

Exactly how many doxastic attitudes does opinionation require us to have? The answer to this questions depends on which claims are supposed to be represented in a given model. This in turn will depend on the modeling language employed, and on how the objects of degree of belief are structured. If the modeling language includes all of the logical combinations of the atomic claims in the modeled agent's natural language or all of the logical combinations of atomic claims towards which the agent has degrees of belief, the opinionation requirement will be very demanding. For instance, if the objects of degrees of belief have a set-theoretic structure (for example, if they are propositions understood as sets of possible worlds), then opinionation would require that an agent who has degrees of belief about $n$ atomic propositions to have degrees of belief toward $2^{2^n}$ truth-functionally distinct propositions.[100] If the doxastic objects have the finer grained structure of sentences in a logical language, the opinionation requirement would require an infinite number of doxastic attitudes, since there are infinitely many sentences in any language containing at least one atomic sentence. We could not, it seems, have so many doxastic attitudes.[101]

[100] That is, for $n$ propositions there are $2^n$ state-descriptions (truth-table lines), and so there are $2^{2^n}$ sets of state-descriptions (sets of truth-table lines), where each such set consists of those state-descriptions that combine to make up a disjunction of state-descriptions (where the empty set corresponds to a contradiction, and the set of all state-descriptions corresponds to a tautology).

[101] Titelbaum (2015) mentions a similar objection to logical omniscience principles. He calls it the "cognitive capacity objection" (2015, p. 255).

Probabilists tend to give two main responses to this worry. First, they remind critics that probabilism is a normative theory.[102] They explain that the epistemic requirements of probabilism characterize epistemic ideals. Their status as such is not diminished by the cognitive short-comings of humans. So the requirement that agents have degrees of belief about every claim is an epistemic ideal despite the fact that human agents couldn't hope to live up to it.

Second, some probabilists offer an additional sort of response. They contend that critics have not merely misapprehended the normative status of probabilism; critics have also misapprehended the aims and method of probabilistic modeling. They explain that probabilistic models should not be undertaken on the global scale, modeling an agent's entire doxastic state at a given time.[103] This project, they suggest, is overly ambitious and complex. Rather, they suggest that each probabilist model should be constructed at a local level to model simplified situations in order to generate normative specific verdicts about how agents ought to invest their degrees of belief. On this approach, each model would specify a small number of atomic expressions. So, the number of truth-functionally distinct expressions included in the model would be relatively small. So, the local approach would not impose an excessively demanding opinionation requirement.

[102] See Christensen (2004, Ch. 6) for detailed discussion of this kind of response to concerns that probabilism is overly normatively idealistic.

[103] Garber (1983) proposes the local modeling strategy, and Titelbaum (2013) defends it at length. See also Levi (1967), Shimony (1970), and Savage (1972) for earlier, less explicit defenses of local modeling.

These responses raise familiar issues about how to interpret probabilist formal models and how to understand their normative significance. The first response saves probabilism from psychological plausibility concerns by a familiar strategy, namely, affirming the status of probabilism as a normative theory, and explaining that normative theories are not beholden to the limitations of actual agents. As we've seen above, it's not obvious that this response is adequate. Probabilists will have to say more about the normative foundations of epistemology before they can deflect this objection by mere reference to the normative nature of probabilism. The second response raises further questions about the purpose of probabilistic modeling. Is the probabilistic modeling framework a mere instrument for generating normative verdicts? Or is it supposed to help explicate some philosophical notion, like *coherence* or *rationality*?

#### 3.1.2.2 Would We Want to Have So Many Doxastic Attitudes?

A related worry concerns whether we, as real agents, would even want to have degrees of belief in such a vast range of claims. Some authors have suggested that having degrees of belief in so many claims would simply clutter our cognitive lives with irrelevancies.[104] A language sufficiently rich to model an agent's doxastic state would likely include many odd claims, including extremely complex logical combinations of unrelated atomic claims. Why would we want attitudes towards extremely complex

---

[104] To my knowledge, Harman (1986) is the first to make this point. Harman's focus is on the normative connection between deductive logic and categorical belief, but his point applies equally well to the normative connection between the probability calculus and degree of belief. See MacFarlane (MS) and Field (2009) for additional discussion.

tautologies or logical combinations of totally unrelated atomic claims? Such attitudes are not epistemically valuable. They would consume our limited mnemonic resources.

Probabilists respond to this worry much like they respond to the previous worry about opinionation. Some probabilists reaffirm that probabilism is a normative theory. They explain that the fact that human minds, if opinionated, would be hopelessly cluttered does not show that opinionation and probabilistic coherence fail to set the appropriate normative standard for degrees of belief. They cite various justifications for probabilism to show how failure to abide by opinionation and the other such normative requirements of probabilism result in terrible consequences.

Other probabilists reiterate the response that opinionation is a demanding cognitive requirement only when viewed from a global modeling perspective. If we adopt a local modeling perspective, the opinionation requirement will not require us to clog our minds with degrees of belief in irrelevant claims. For the modeling languages employed in probabilist models will be specifically designed to yield limited normative verdicts. These models will employ very limited modeling languages, and local modelers will not focus on generating verdicts about irrelevant logical combinations of claims.

Again, these responses raise questions about what the proper normative interpretation of probabilism should be, and about how we should understand the aims and methods of formal probabilist modeling. In addition to raising questions about whether probabilism is an evaluative theory, the first response also raises the question about what value of rational degrees of belief are supposed to have. Is this value determined by practical interests, so that only degrees of belief relevant to one's interests are valuable? Or are all rational degrees of belief valuable in some wider sense, even if they are not relevant to one's practical interests? And as before, the second response raises the question of whether probabilist models should be local or global, but it makes especially perspicuous the question of whether the normative interest of probabilism is in its local or global application.

#### 3.1.2.3 Too Hard to Compute?

Another related concern is that opinionation, in conjunction with the other probabilist requirements, requires us to maintain probabilistic coherence for a huge number of doxastic attitudes. As the previous worries contend, humans are cognitively limited agents. It doesn't seem like we could hold all of the attitudes opinionation requires, and it seems like it would be cognitively harmful for us to try. Relatedly, our cognitive limitations would seem to prevent us from computing the coherence of so many degrees of belief. Probabilistic coherence (a.k.a. probabilistic consistency with the probability calculus) on a language for sentential logic (or on the language of Boolean combinations of sets) is decidable (Fitelson 2008). But the decision problem for probabilistic coherence is more computationally complex than the decision problem for the logical consistency of sets of sentential logic sentences. In some realistic cases, the sun would die long before an agent could compute the coherence of all of their probabilistic degrees of belief.

Probabilist responses to this concern are similar once again to those just presented with regard to the other worries about opinionation. Probabilists might refer once more to the fact that probabilism is a normative theory. The computational limitations of humans do not impose constraints on its normative requirements. It is still epistemically good or obligatory to maintain probabilistic coherence among an opinionated degree of belief set, even though a human could not do so (just as deductive logical consistency may be the appropriate normative standard for categorical belief, but computationally beyond reach in many realistic cases). Or probabilists might lean on the local modeling approach again, noting that probabilist models do not impose immodest computational demands when approached from a local perspective.

As before these responses raise questions about the normative significance of probabilist requirements, and about how to conceive of the probabilist modeling enterprise. This worry poses an extreme example of the objection that probabilism should be constrained by the abilities of human agents. By toeing the normative line in response to this worry, the probabilist seems to admit that it is good (or perhaps, obligatory) to maintain a coherent opinionated degree of belief set, even though in some realistic cases computing the coherence of such a set could take longer than humans have existed.

This response seems to suggest that the real cognitive abilities of humans place absolutely no constraints on the normative requirements of probabilism. This raises deep questions about the nature and source of probabilism's normative force. The second response also provokes worries about the normative significance of the local modeling approach. As usually presented, the primary normative claim of probabilism is its requirement that agents should maintain probabilistically coherent degrees of belief. If this requirement can be watered down to a sufficiently local level, probabilism seems to lose much of its normative force. It seems to reduce to the claim that modeling belief strengths with probabilities is sometimes useful for small, locally tractable problems. Who would have doubted that?

### 3.1.3 Logical Omniscience

As we saw in Chapter 2, probabilism seems to impose a logical omniscience requirement on agents. It requires agents to be representable with probabilistic credence functions. And Kolmogorov's second probability axiom (normality)[105] requires all logical truths to have probability 1. If we translate the normality axiom into a degree of belief norm by applying the probabilist scheme for representing degrees of belief in reverse, this seems to yield the requirement that agents must be certain of all logical truths. (On the set-theoretic version of the axioms, the equivalent requirement is that all set-theoretic expressions for the set $\Omega$, no matter how complex, be recognized to equal $\Omega$, and thus be assigned probability 1.)

However, this is not the only logical omniscience requirement that probabilism seems to impose. For, in conjunction with the other axioms, the normality axiom entails theorems which translate into norms that impose additional logical omniscience constraints. For example, one such norm holds that we should be at least as confident in each claim as we are in the claims that entail it. Another holds that we should be certain that all contradictions are false.

[105] For all $A \in \mathcal{L}$, if $\vDash A$, then $pr(A) = 1$.

Additionally, the opinionation requirement can itself be construed as a kind of logical omniscience requirement. For it holds that we should have attitudes toward all of the logical (set-theoretic) combinations of the basic claims towards which we have attitudes. Awareness of these logical (set-theoretic) combinations is itself a form of logical omniscience.[106]

Critics of probabilism raise worries about the logical omniscience requirement similar to those they raise about the opinionation requirement, as well as a couple of others. I present these in sequence below.

#### 3.1.3.1 The Problems

##### 3.1.3.1.1 Excessive Demandingness

The most common objection to the logical omniscience requirement is that it is too demanding. Critics object that the logical omniscience requirement is too demanding for two reasons.[107] First, there's what Titelbaum (2015, p. 255) calls the "cognitive capacity objection." As I just explained, the logical omniscience requirement seems to demand that we should be certain in the truth of all logical truths, certain of the falsehood of all contradictions, and at least as confident of each claim as we are in claims that entail it. Having so many doxastic attitudes, the objection holds, is beyond our cognitive capacity. For, even in sentential logic there are infinitely many logical truths and falsehoods, and infinitely many entailment relations among claims. Humans

---

[106] Earman (1992, p. 122) calls this form of logical omniscience "LO2." He refers to the demand for certainty in all logical truths as "LO1" (1992, p. 121).

[107] See again Christensen (2004, Ch. 6). See Cherniak (1986) for a discussion of the demandingness of the deductive consistency norm for categorical belief.

couldn’t possibly have infinitely many doxastic attitudes. We simply could not maintain the sheer number of degrees of belief required. Our mnemonic abilities fall far short of this standard.

Second, there’s what Titelbaum (2015, p. 255) calls the “cognitive reach” objection: some logical truths, logical falsehoods, and entailment relations are too complex or obscure for us to recognize them as such. Our computational limitations also prevent us meeting this standard.

These objections suggests that probabilism is an inappropriate norm, because it makes demands that real agents couldn’t hope to meet. In this regard probabilism is on a par with a logical consistency norm for categorical belief, which would require us to believe all logical truths and to disbelieve all contradictions.[108]

**3.1.3.1.2 Cognitive Clutter**

Relatedly, critics also raise cognitive clutter concerns about the logical omniscience requirement. Even if we could be logically omniscient, it’s not clear that all of this logical awareness would be of any value to us. We might worry that so many degrees of belief (probabilistic or not) would clutter our minds with irrelevancies unimportant to our pragmatic and epistemic ends. And since our cognitive limitations prevent us from having so many degrees of belief, these cognitive clutter concerns are even more pressing. For many logical truths and falsehoods are extremely obscure and distant from

[108] Christensen (2004, p. 151) also makes this point.

our practical lives. To cite a recent example from the literature,[109] there are logical truths (consequences of the axioms for the real numbers) that concern the trillionth digit in the decimal expansion of $\pi$. Even if certainty in logical truths were intrinsically valuable, surely we would be better off having degrees of belief about matters more relevant to our practical interests than about such obscure logical truths.

**3.1.3.1.3 Rational Uncertainty in Logical Truths**

Critics also note that the logical omniscience requirement seems to prevent probabilists from accommodating rational uncertainty in logical truths. It seems like we may adopt rational, yet uncertain (below credence 1) degrees of belief in claims that turn out to be logical truths. But probabilism seems to require us to be certain of all logical truths. And if we must be certain of all logical truths, then we can't very well be *un*certain of any.

Titelbaum suggests the following example, where his uncertainty in a logical truth seems to be warranted (2013, p. 106). Titelbaum explains that Talbott (2005) claims that the trillionth digit in the decimal expansion of $\pi$ is 2 on the basis of having consulted Google on the matter. Before reading Talbott's article, Titelbaum was uncertain about whether or not the trillionth digit of π is 2—his degree of belief could have been modeled with a credence .1 on the basis that he thinks that the occurrence of each digit (from '0' through '9') is equi-probable because π is a "normal" number.[110] After reading

[109] See Titelbaum (2013, p. 106 ff.) who cites Talbott (2005). Similar examples can also be found in Savage (1967) and Hacking (1967).

[110] That is, no digit 0–9 is more likely than any other to occur at any point in its decimal expansion.

Talbott's article, Titelbaum's confidence that the trillionth digit of $\pi$ is 2 increased greatly, though he was still not totally certain. After all, Talbott (and Google) could be mistaken. Further corroborating evidence could raise Titelbaum's confidence even more. Titelbaum's response to his new information, and his high, but uncertain credence that the trillionth decimal of π is 2 seem to be rational given his evidence. His reliance on Talbott and Google seems to support high confidence, but not certainty. But probabilism, in light of the logical omniscience requirement, seems to yield the verdict that Titelbaum's credence is irrational. He has an uncertain degree of belief in a logical truth. Thus, *prima facie*, the logical omniscience requirement leads probabilists to make the wrong judgment in some cases about which degrees of belief are rational. As Titelbaum (2013, p. 106–107) notes, this problem with the logical omniscience requirement has not been well-acknowledged, less so than the preceding worries.

**3.1.3.1.4 Rational Logical Learning**

The logical omniscience requirement leads to a related worry concerning rational logical learning. Considering the preceding example once again, it seems like Titelbaum's revision of his credence in light of his new evidence is rational. Titelbaum had a reasonable initial degree of belief based on his evidence, and he revised his degree of belief reasonably in light of new evidence after reading Talbott's article. But probabilism seems to require the verdict that Titelbaum violated the standard of ideal rationality when he initially deemed it equiprobable than any digit 0-9 was the trillionth digit of π, and that Titelbaum was still irrational when he became more confident, but still uncertain, that the trillionth digit is 2. Any degree of belief short of certainty (in whatever the actual value of the digit happens to be) violates ideal rationality. Probabilism cannot accommodate Titelbaum's apparent logical learning. In general, the logical omniscience requirement seems to prevent probabilists from allowing that one can rationally gain confidence in logical truths. Even upon completing a proof, one might not yet assign certainty to the result. After all, logicians make mistakes. But probabilism calls irrational any revision of degrees of belief in logical truths that yields a result short of certainty. Rational logical learning can only occur when one remedies a flaw in one's degrees of belief about logic by gaining certainty in logical truths.

#### 3.1.3.2 Responses

##### 3.1.3.2.1 Changing the Scope of Possibility—Hacking

Probabilists offer several different responses to these worries. Hacking (1967) suggests avoiding these problems by characterizing a notion of "personal possibility." A sentence is personally possible for an agent if the agent can understand it and the agent doesn't know it to be false (Hacking 1967, p. 318).[111] Hacking suggests limiting the probabilist modeling language to the set of sentences the modeled agent can understand, and modifying the probability axioms also so that they only demand agents to maintain coherence among the sentences they understand that aren't ruled out by the agent's knowledge. Thus, when it comes to logical truths, Hacking suggests modifying the logical omniscience requirement to the requirement that agents must be

---

[111] Hacking assumes that knowledge is not necessarily closed under entailment (1967, p. 319). For example, he says that one could know that $A \rightarrow B$ and that $A$, yet nonetheless $\neg B$ could be possible for the agent.

certain of all of the logical truths that they understand.[112] This is a sort of fore-runner to the local modeling approach. Hacking sees this approach as characterizing a notion of "personal probability"[113] that fits within a normative hierarchy, with the traditional notion of probabilistic coherence sitting high above personal probabilistic coherence, but just below total omniscience—logical and factual (1967, pp. 320-321).

This response makes significant headway in diminishing demandingness and cognitive clutter concerns about the logical omniscience requirement. It is much more plausible that we can have degrees of belief in all of the logical truths we can understand than in *all* of the logical truths. And so the class of truths we can understand threatens much less cognitive clutter. The computational demands are also greatly diminished. At the very least we will be able to compute enough to recognize each logical truth as such, though we may not initially recognize it as such, even if the computation of the coherence of these truths with all of our other degrees of belief remains beyond our grasp. But Hacking's response does not extinguish the demands that the logical omniscience requirement places on us. Being certain of every logical truth we can understand will require significant cognitive resources at the cost of other perhaps more valuable pursuits.

Thus, as Hacking admits, this response is still subject to concerns about demandingness and cognitive clutter. And thus it is not immune to concerns that it is

---

[112] Since knowledge is factive, an agent couldn't know a logical truth to be false.

[113] Not to be confused with the notion of "personalist probability," which is another name for subjective probability or degree of belief.

normatively inappropriate. But this response also provokes several other questions. For it seems to soften the demands of the logical omniscience requirement at the cost of being normatively too weak.[114] There are logical truths for which we might want to hold agents normatively accountable even if they do not understand these claims because we think they should understand. For example, we might think an agent ought to be certain of a simple tautology of the form $A \vee \neg A$ (where ⌜ $A$ ⌝ is an atomic sentence, let's say), even if the agent doesn't understand it. This brings us back to the issue of what level of normative strength is appropriate to probabilism. Hacking suggests a hierarchy of epistemic norms associated with different purposes. He suggests that both the highest ideals on this hierarchy and the more realistic standard are both essential to epistemic theorizing. This is an idea I will return to in Chapter 5.

#### 3.1.3.2.2 Local Modeling—Garber

Garber (1983) offers a different approach that also requires limiting the modeling language. In general, Garber advocates a local modeling approach. With respect to modeling logical awareness, Garber advocates limiting the modeling language to a sentential language. On this approach, all non-sentential logical truths are treated as extra-systematic constraints to be specified by the modeler in particular cases. Thus, the probabilist need not require agents to be certain of all logical truths at all times, but only of the logical truths of sentential logic, and, perhaps, particular further logical truths as specified by the modeler.

[114]Titelbaum (2013, p. 110) attributes this point to Eells (1985).

Like Hacking's response, this response greatly diminishes the force of the demandingness and cognitive clutter concerns. Limiting the omniscience requirement to the sentential logical truths (and some higher-level truths specific to particular circumstances) removes a significant class of truths from the set for which agents seem to be accountable. But, as before, this response still seems to require agents to have more degrees of belief than they could have. There are still many, many sentential logical truths. And many sentential logical truths are still far too complex for human agents to be expected to recognize them as logical truths. Additionally, as before, this response does not obviously assuage concerns about logical uncertainty and logical learning. The extrasystematic constraints imposed by modelers in particular situations might help to weaken these concerns. But Garber's account still seems to make rational logical uncertainty and rational logical learning of sentential logical truths impossible.

Garber's response seems to be superior to Hacking's in that it is normatively stronger in a sense. For Garber's approach requires agents to be certain of all truth-functional truths represented in the modeling language, whereas Hacking's approach merely demands certainty in the logical truths the agent can understand. But there may be obvious truth-functional truths the agent doesn't understand, though they should, as I explain above. Garber's approach demands certainty in these truths, but Hacking's does not. So, Hacking's approach permits a form of logical obtuseness[115] by permitting uncertainty in logical truths the modeled agent doesn't understand. But Garber's view

[115] The term "logical obtuseness" is borrowed from MacFarlane (MS).

also seems to be too strong in another sense.[116] It still imposes very high cognitive demands by requiring truth-functional omniscience. Once again, this raises the issue of how strong the normative demands of probabilism ought to be interpreted as being.

**3.1.3.2.3 Coherent Extendability—Zynda**

Zynda (1996) gives a different response. Rather than proposing ways to limit the class of truths for which agents ought to be held accountable, Zynda suggests a reinterpretation of the normative import of probabilism. He explains that the probabilist norm, including the logical omniscience requirement, does not require agents to adopt probabilistic degrees of belief in particular claims. Probabilism is simply a coherence requirement. It requires only that *if* agents invest credence in particular claims, then those degrees of belief must be modelable with credence functions that obey the probability calculus. Thus, probabilism does not require agents to be certain of all logical truths and certain of the falsehood of all contradictions. It merely requires that, *if* an agent has a degree of belief about a logical truth, then they should be certain of it. In other words, agents' degrees of belief should be coherently extendable to credence sets that assign certainty to all of the logical truths.

Thus, Zynda's response seems to largely avoid concerns about demandingness and cognitive clutter. Probabilism doesn't require us to be certain of all logical truths. It doesn't place the attendant enormous cognitive demands on us. It merely requires that *if* we adopt degrees of belief in logical truths, that these degrees of belief be certainties.

[116] Again, Titelbaum (2013, p. 110) attributes this point to Eells (1985).

Thus, you might say that it requires logical *infallibility* rather than logical *omniscience*. It doesn't require you to be certain of all logical truths. Instead, it requires that if you adopt an attitude toward a logical truth, it must be the attitude of certainty.

This isn't to say that probabilism, as Zynda understands it, doesn't place any strong cognitive demands on agents. Understood this way, probabilism will still require agents to maintain coherent credence sets—credence sets that are extendable to opinionated logically omniscient sets at that. This is still a very demanding requirement. It requires agents to have excellent memory and to be excellent computers. Many agents will encounter some very complex logical truths, and probabilism will still require them to be sure of these truths and to incorporate these certainties coherently into their total credence sets. But then, it is no surprise that probabilism is a demanding requirement. For as Zynda notes, it is intended as a normative ideal. Furthermore, as Titelbaum notes, this kind of response does not resolve the issues concerning rational logical uncertainty and logical learning. On this view, probabilism still requires agents to be certain of the logical truths they do invest credence in, and to hold degrees of belief that are coherent with certainty in these truths even if the agent doesn't actually consider such truths. So, this version of probabilism still yields the verdict that any agent who has a degree of belief less than certainty for a logical truth violates the standards of ideal rationality.

Zynda's response raises a couple of issues about how to interpret probabilism's normative significance. First, it raises the issue of why logical infallibility is an appropriate normative requirement, but logical omniscience is not. Zynda's response dodges many concerns about demandingness and cognitive clutter. But it isn't clear that this is a good thing. We need to appeal to firm principles that place constraints on normative theorizing. Second, it raises the issue of what kind of verdicts we want probabilism to produce. On Zynda's view, probabilism gives verdicts only about incoherence. It never requires agents to adopt some specific degrees of belief for a claim. Thus, like Hacking's approach, it also permits a form of logical obtuseness. But it seems sometimes appropriate to judge an agent to have violated the ideals of rationality for failing to adopt some specific degree of belief in a simple, epistemically accessible claim. Zynda's approach fails to do so. Here again, we need to appeal to principles of normative epistemic modeling to adjudicate such questions.

**3.1.3.2.4 Titelbaum**

Titelbaum's (2013, Ch. 5) response to the logical omniscience worries incorporates elements of each of the foregoing accounts. Like Hacking and Garber, Titelbaum advocates a local modeling approach.[117] He takes the point of probabilism to be the delivery of normative verdicts about agents' degrees of belief in particular situations. To arrive at these verdicts, he thinks formal modelers should specify simplified versions of situations, and subject these to small, local probabilist models that yield the appropriate normative verdicts.

[117] See Titelbaum (2013, Chs. 2-4) for a detailed account of his modeling methodology. See Titelbaum (2013, p. 33, n. 2) for an explicit connection between this methodology and Garber's (1983) notion of local modeling. And see Titelbaum (2013, pp. 108-110) for an explanation of how his framework accommodates logical omniscience problems and how it relates to Hacking's and Garber's approaches.

Like Zynda, Titelbaum also advocates limiting the normative import of probabilism to judgments of incoherence.[118] So, on his view, probabilism never requires an agent to adopt a particular degree of belief in a particular claim. Rather, Titelbaum's version of probabilism yields negative evaluations of agents whose degrees of belief are either explicitly incoherent or which cannot be coherently supplemented with additional degrees of belief.

Titelbaum (2013, pp. 108-110) also employs a further modeling technique that insulates his formal probabilist modeling framework from logical omniscience worries. Titelbaum makes a version of the finite additivity axiom the only essential constraint of his formal modeling system common to every model (2013, p. 45). He accounts for the content of the other standard probability axioms in terms of standard extrasystematic certainty conditions required by his formal models. Thus, in most cases, Titelbaum's formal models will include a requirement akin to logical infallibility—interpreted as an injunction to maintain degrees of belief that are extendable to a set that is modelable with a probabilistic credence function—rather than demanding full-blown logical omniscience.

However, Titelbaum acknowledges that in some cases in which agents entertain degrees of belief in logical truths, the standard interpretation of Titelbaum's probabilism may not be appropriate due to logical omniscience concerns (2013, pp. 108). In these cases, Titelbaum advocates modifying his version of probabilism to avoid inappropriate

[118] See Titelbaum's discussion of his "Evaluative Rule" (2013, pp. 56-59).

normative verdicts—for instance the verdict that it is incoherent for an agent to be uncertain that 2 is the trillionth decimal of π. Titelbaum makes two suggestions about how to make the appropriate modifications. First, he suggests that in cases where logical omniscience worries arise, rather than requiring agents to be certain of all logical truths, we should require agents to be certain of all obvious logical entailments (2013, p. 110). Titelbaum also considers a second option, namely that in some cases it might be appropriate to abandon any firm extrasystematic constraints on degrees of belief in logical truths (2013, p. 109-110). In some cases, he suggests, it should be left entirely up to the modelers to decide which logical truths are appropriate as required certainties for agents.

Like Zynda's view, Titelbaum's view largely avoids concerns about demandingness and cognitive clutter. But Titelbaum's view also offers a response to concerns about modeling rational logical uncertainty and logical learning. In the example of Titelbaum's degrees of belief about the trillionth digit of π, neither of Titelbaum's proposed modeling responses will judge him to have violated ideal rationality. For truths about the trillionth digit of π are far from obvious, and modelers worth their salt would not judge such uncertainty irrational. As Titelbaum (2013, p. 110) notes, this requirement is stronger than Hacking's requirement that agents be sure of all of the logical truths they understand, but weaker than Garber's requirement that agents be sure of all sentential logical truths. It allows us to judge agents negatively for their incoherent degrees of belief (or commitments to degrees of belief) in claims they do not understand, but it does not require us to be certain of any huge class of logical truths.

Titelbaum's response raises some of the same questions as Zynda's concerning what kind of normative verdicts we want probabilism to produce and what kind of normative force we want these verdicts to have. But Titelbaum's response also raises questions about how we should apply formal models to formulate doxastic norms and generate normative verdicts about doxastic attitudes. It raises the question of whether we want probabilism to provide global norms that apply in all cases, or whether we want to generate local norms that apply only to simplified versions of particular cases. As before, to answer these questions adequately, we will need principled views about the purpose of formal modeling and the intended normative force of the probabilist norm.

### 3.2 Digestion of Dialectical Themes

Now that I have presented this portion of the dialectic between probabilists and their critics, I will draw out some of the major themes from their exchange of objections and replies. My purpose will be to show how this dialectic raises deep questions about probabilism's foundations, and how it begins to reveal various probabilist answers to these questions.

#### 3.2.1 Themes from the Objections

Above I presented objections to three aspects of probabilism: its formal representation of degrees of belief, its opinionation requirement, and its logical omniscience requirement. The objections to these aspects of probabilism fall into three main kinds: descriptive adequacy objections, demandingness objections, and value objections.

Most of the objections to probabilism's formal representation of degrees of belief *via* probability functions considered above are challenges to its descriptive adequacy. These objections challenge probabilism on the grounds that its characterization of degrees of belief via the formal probabilistic representation is inaccurate and inadequate. Some of these objections are based on misapprehensions of the view. For example, the *no numbers in the head* objection conflates probabilism's credence functions with the degrees of belief they are designed to represent, which need not be numerical. And most versions of the objection that actual agents don't have probabilistically coherent degrees of belief misapprehend probabilism as a purely descriptive theory, mistaking the probabilist norm for a descriptive claim. There are also other descriptive adequacy objections, however, that hit closer to the mark. For probabilism is not an entirely normative theory: it posits the existence of degrees of belief within the minds of human agents (in some sense of "within"), and models them with credence functions. As such these aspects of probabilism are clearly subject to descriptive adequacy constraints. And, as we saw above, there are good reasons to question the representational faithfulness of probabilism's precise numerical representation of degrees of belief.

These descriptive adequacy objections raise various methodological questions about probabilism's formal representation scheme. For example, they raise various questions about how realistically credence functions represent degrees of belief: How exactly are credence functions meant to model degrees of belief? What is the underlying characterization of degrees of belief, and how do the mathematical properties of credence functions correspond to the properties of degrees of belief? The descriptive adequacy objections also raise the issue of how much methodological idealization—how much descriptive inaccuracy—is tolerable in the representation of degrees of belief *via* credence functions.

While the purely descriptive aspects of probabilism are clearly subject to descriptive adequacy conditions, there are also reasons to think its normative aspects should be constrained by considerations about what human beings and our mental states are really like. For, if probabilism makes normative demands or recommendations that humans can't live up to, it seems that its normative interest and promise are diminished. Many of the objections to probabilism considered above are objections of this kind—demandingness objections. These objections hold that the probabilist norm is implausible because it is too demanding. Thus, these objections seem to assume a kind of "ought"-implies-"can" principle. Some versions of the psychological objection to the probabilist representation scheme and some of the objections to the logical omniscience and opinionation requirements are demandingness objections. These demandingness objections clearly raise questions about whether and how the normative aspects of probabilism ought to be constrained by facts about human abilities and psychology. However, the demandingness objections also raise questions about the probabilist norm's normative force, and the purpose of normative probabilist modeling. The probabilist responses to these objections raise further questions in this vein.

In addition to challenging probabilism's descriptive adequacy and demandingness, some of the objections surveyed above also challenge the value of achieving probabilistic coherence among our degrees of belief. For example, the evidential objection to the probabilist representation of degrees of belief is a challenge to the value of having precise degrees of belief that admit of point-valued numerical representation. Because our evidence is sometimes imprecise, incomplete, or equivocal, the objection holds, representation via numerically precise credences is not always desirable. This objection raises the question of how much methodological idealization is tolerable from a normative standpoint. What benefits does the computational and representational tractability of point-valued representations of degrees of belief provide, and do these benefits outweigh any tendency to produce inappropriate verdicts in cases where the agent's actual degrees of belief are incomplete, imprecise, or equivocal given the available evidence?

The cognitive clutter objections to the opinionation and logical omniscience requirements can also be construed as value objections. For, they challenge the value of having opinionated and logically omniscient degrees of belief. One version of the clutter objection is a kind of mix between a demandingness objection and a pure value objection: given that we have limited mnemonic and computational abilities, it would be a poor use of our cognitive resources to pursue opinionation and certainty in all logical truths. This version of the objection challenges the value of opinionation and logical omniscience given our cognitive limitations. But the clutter objections are not necessarily based on our cognitive limitations. Stronger versions of these objections question the value of opinionation and logical omniscience even on the assumption that we could be opinionated or logically omniscient. Even on that assumption, according to the stronger versions of the objections, it wouldn't be valuable for us to clutter our minds with an infinite number of attitudes—an infinite number of which will never have any practical impact on our lives. This stronger version of the clutter objection is especially challenging when applied to the opinionation requirement. For, while it seems like it would be useful to be certain of all of the logical truths, if one could be, it's not clear why one would want to have degrees of belief in all of the logical combinations of claims.

Like the demandingness objections, the challenges to probabilism's value posed by this form of clutter objection raise questions about the purpose of probabilist modeling and the normative force of the probabilist norm. But they also raise questions about the justification of the probabilist norm and the source of its normative force. Why should we value opinionation and logical omniscience, and what kinds of value do they have?

### 3.2.2 Themes from the Probabilist Responses

The probabilist responses to the objections surveyed above break down into four main kinds in terms of two distinctions. Some of the responses are what I call concessive strategies, while others are dogmatic. As you might expect, the concessive strategies concede to the objection and modify probabilism in order to accommodate the objection. Dogmatic strategies, on the other hand, either deny the force of the objection or they hold that, while the objection does have force, it does not merit modifying probabilism. We can distinguish the responses further according to whether they concern methodological or normative issues.

Some of the probabilist responses to descriptive adequacy objections are methodologically concessive. For example, when adopted as ways to make the probabilist representation of degrees of belief more psychologically realistic or more descriptively adequate, imprecise or qualitative forms of probabilism are methodological concessions. They amount to admissions that the point-valued numerical representation of degrees of belief is too precise and therefore descriptively inadequate.

Other probabilist responses to descriptive adequacy objections are methodologically dogmatic. They admit that the point-valued representation of degrees of belief is often descriptively inaccurate, but they maintain that the increased computational tractability of working with point-valued credence functions rather than sets of such functions or, alternatively, comparative confidence relations is worth the cost in terms of descriptive accuracy.

These responses begin to answer methodological questions about the probabilist representation scheme raised by the descriptive adequacy objections. The concessive approaches characterize degrees of belief to allow imprecise attitudes, and they admit that computational tractability and representation simplicity are not worth the cost of descriptive inaccuracy. The dogmatic approaches, on the other hand, take the opposite position. While these different approaches begin to answer the questions raised above—and, thus, they begin to fill in probabilism's methodological foundations—they still leave many questions unanswered. In particular, they do not provide complete or explicit answers to the question of exactly how the mathematical features of the probabilist formalism model the properties of degrees of belief. They also leave open the question of whether the concessive approaches or the dogmatic approaches are superior.

Most of the probabilist responses to demandingness and value objections are normatively concessive. For example, while the local modeling approach is intended primarily to make probabilist modeling of particular cases more tractable, it is also intended to ameliorate demandingness concerns about the opinionation and logical omniscience requirements. For, on the local modeling approach, a given probabilist model is intended only to represent a very small subset of an agent's degrees of belief. So, the opinionation and logical omniscience requirement, while they will still require an infinite number degrees of belief,[119] will require degrees of belief in only a relatively small (finite) number of logically distinct propositions.

In addition to conceding somewhat to demandingness and value objections, the local modeling approach takes a very clear stand on the purpose of probabilist modeling: namely, the primary purpose is to generate verdicts about how to apportion our confidence in particular cases. Thus, the local modeling response makes a very clear contribution to the specification of probabilism's foundations.

Adopting a set-theoretic approach to modeling the objects of degrees of belief—in particular, approaches on which the doxastic objects are epistemic, doxastic, or personal possibilities—can also be seen as a concession to demandingness and value concerns.

---

[119] Here I'm assuming that the objects of degree of belief are as finely-grained as sentences in a formal language.

Such approaches appear to weaken the demands of the logical omniscience and opinionation requirements due to the extensionality of sets. After all, for a finite number of atomic doxastic objects, opinionation and logical omniscience will require merely finitely many total doxastic states on a set-theoretic approach. So, adopting this approach seems to be at least a minor concession to claims that logical omniscience and opinionation are too demanding.

Adopting the extendability form of the probabilist norm is also a concessive strategy. In particular, this form of the norm concedes to demandingness and value objections to the opinionation and logical omniscience requirements. For it does not require opinionation in any form, and it does not require a full-blown form of logical omniscience. It requires the somewhat weaker demand of logical infallibility. Thus, it seems to grant that the opinionation and logical omniscience requirements are either too demanding or lack sufficient value. This response begins to specify the normative connections between the probability calculus and degrees of belief.

Additionally, while modifications to the probabilist formalism can be methodological concessions, they can also be normative concessions—concessions that the precision of probabilism's point-valued representation of degrees of belief is excessive and of dubious value. Both popular modifications to the probabilist formalism—those that represent degrees of belief with sets of credence functions and those that represent them with a confidence ordering relation—allow imprecise, incomplete, or equivocal degrees of belief in the face of imprecise, incomplete, or equivocal evidence. Thus, in response to the question of whether some normative inappropriateness is worth the cost of improved computational tractability and representational simplicity, these responses answer “no.”

While most of the probabilist responses to demandingness and value objections are concessive, the evaluative reinterpretation response is most naturally viewed as dogmatic. According to this response, the probabilist norm has evaluative normative force, and as such it is not subject to “ought”-implies-“can” concerns. Like the extendability response, this response also begins to specify the precise normative connections between the probability calculus and degrees of belief.

### 3.3 Conclusion: The Need for Unified Foundations

In this chapter, I have presented central issues in the dialectic between probabilists and their critics, and I have drawn out some of the major themes from this exchange. I presented three main sources of objections to probabilism, and I presented several of the major responses to these objections. I showed how the objections I surveyed fall into three main kinds (descriptive adequacy, demandingness, and value), and how the probabilist responses tend to be concessive or dogmatic in the face of these concerns. I also showed how this dialectic raises various questions about probabilism’s methodological and normative foundations, and how probabilist responses to objections begin to reveal different answers to these foundational questions.

I want to conclude this chapter with an argument that probabilists must articulate unified foundations for their view. My basis for this conclusion is simple. As it is characterized in Chapter 2, probabilism is subject to several powerful objections. These objections raise various important questions about probabilism’s foundations, its content, and its aims. While probabilists have various responses to the objections, responses which reveal answers to foundational questions, the various responses provide different kinds of answers. Indeed, in some cases the responses are at odds—and, thus, appear not to be co-tenable. In some cases this incoherence is blatantly obvious. For example, it wouldn't make sense to advocate modifying the probabilist representation scheme to make it more descriptively adequate while also holding that such modifications are unnecessary or not worth the cost in terms of computational tractability and representational simplicity. Indeed, in general, it doesn't make much sense to endorse any of the opposing concessive and dogmatic responses to the same objection.

However, in some cases, the incoherence of different responses is subtle. For example, recently it has become popular among probabilists to advocate both the extendability form of the probabilist norm, and an evaluative interpretation of the norm. According to this view, the probabilist norm says that one's degrees of belief ought to be coherently extendable to a richer—opinionated and logically omniscient—degree of belief set that can be faithfully modeled with a probability function. This version of probabilism doesn't require opinionation, and it requires logical infallibility rather than logical omniscience. Additionally, the force of the "ought" in the probabilist norm is that of evaluation rather than obligation or responsibility. It is *good* for one's degrees of belief to be extendable to probabilistic coherence.

At first glance, these responses seem perfectly coherent, for both of them seem to be oriented at avoiding demandingness objections to the opinionation and logical omniscience requirements. The problem, however, is that the two responses set out to avoid these concerns in ways that seem to be at odds. The extendability approach modifies the probabilist norm to weaken its demands. Thus, it concedes to the force of demandingness objections. The evaluative reinterpretation response, however, is dogmatic in the face of demandingness concerns. It holds that the probabilist norm is not subject to demandingness concerns because it is an evaluative principle, and evaluative principles like the probabilist norm are not subject to “ought”-implies-“can” type constraints. These two responses seem to be at cross purposes. The extendability form of the norm seems to be intended to bring satisfaction of the norm within the reach of real agents, while the evaluative reinterpretation of the norm seems to be intended to side-step demandingness concerns by specifying an ideal for degrees of belief that does not take the abilities of real agents into account. In the end, I think these responses are compatible. However, I think it is unclear—at least initially—why probabilists would want to endorse both responses.

Thus, I contend that probabilists should provide unified foundations for their view in order to make the content and aims of the view clear, to coherently rebut the objections presented in §3.1, and to avoid using the view to derive illicit conclusions on the basis of a misunderstanding of the view’s content or aims. In the remainder of this dissertation, the task I have set for myself is to begin the project of spelling out unified foundations for probabilism. I start with probabilism’s methodological foundations in Chapter 4, where I provide an account of probabilism’s aims, its formalism, and its modeling methodology. In Chapter 5, I provide an account of probabilism’s normative foundations. In particular, I defend an account of the translation of the axioms and theorems of the probability calculus into degree of belief norms and an account of the justification of these principles.

# 4 Methodological Foundations

At the end of Chapter 3, I argued that probabilists need to provide a unified account of probabilism's foundations in order to clarify its aims and consequences, and in order to rebut common objections. The purpose of this chapter is to provide an account of probabilism's methodological foundations. I start with an account of the nature and aims of probabilist modeling. Along the way I provide an account of the nature and aims of formal modeling in general, and probabilist modeling in particular.

After I characterize the nature and aims of probabilist modeling, I begin to present and defend the particular probabilist modeling framework that I think makes the best sense of probabilism, which I call the Comparative Confidence Framework—CCF for short. There are three key aspects of CCF's modeling methodology that jointly distinguish it from extant alternatives in the literature. First, CCF is a probabilist formal modeling framework[120] rather than the mere statement of the probabilist norm in formal terms. Second, CCF employs a logical rather than set-theoretic formal system. Third, CCF employs a fundamentally qualitative and comparative—rather than quantitative—representation of degrees of belief. I will carry on the project of articulating CCF in Chapter 5, where I will characterize CCF's normative foundations.

## 4.1 The Nature and Aims of Probabilist Modeling

As I just explained, the primary purpose of this chapter is to provide an account of probabilism's methodological foundations. However, probabilism's methodology

[120] In the sense of Titelbaum (2013), to be described below.

depends crucially on its aims. So, before I specify my version of probabilist methodology, my aim in this section is to defend an account of the nature and aims of probabilist modeling. Since probabilism is a species of formal modeling in general, I'll begin with an account of what formal modeling is and why we engage in it. Then I'll provide a detailed account of what goes into a formal modeling framework. Finally, I'll provide an account of the specific aims of probabilist modeling.

### 4.1.1 The Nature of Formal Modeling in General

Formal modeling is the use of formal tools—like logic and mathematics—to represent and gain insights about features of the world. Somewhat ironically, it's difficult to give a precise account of formality in this sense.[121] So, I'll rely on a rough characterization according to which the tools of logic and mathematics are formal in the sense that we can use the structures they describe to provide abstract, schematic representations of features of the world.[122] So, in formal modeling, we recognize similarities between formal structures and features of the world. We use these formal structures to represent and gain insights about the target phenomena that interest us.

Formal modeling figures prominently in the methodologies of the sciences and somewhat less prominently in philosophy. Scientists rely on descriptive mathematical models of physical, chemical, biological, psychological, and sociological phenomena in order to recognize important features of these phenomena and to better understand the relationships between their scientific concepts. Likewise, philosophers employ

[121] For an excellent discussion, see MacFarlane (2000).

[122] This rough characterization of formal methods is based on that in Hájek (MSc).

logical and mathematical models to better understand and explicate their philosophical concepts and arguments, and to apply these concepts and arguments to particular cases.

In the case of probabilism, we use the formal structures of credence functions and the probability calculus to gain insights about degrees of belief and degree of belief norms. We recognize the similarity between the structure of the real line between 0 and 1 and the gradational structure of degrees of belief, and we use the structure imposed by the probability axioms to characterize a notion of coherence for degrees of belief.

The models we generate using formal tools can be descriptive or normative, and empirical or *a priori*. We're all familiar with descriptive, empirical formal modeling in the context of science. Scientists of many different stripes use mathematical tools to aid in their descriptions of various physical, chemical, biological, and psychological, and social phenomena. However, we can also use formal tools to model phenomena outside the context of empirical science. Philosophers commonly use logic and mathematics to model philosophical concepts like knowledge, justification, and necessity. And normative uses of formal modeling are also common. Examples of this sort of modeling are most common in the study of reasoning and decision-making. But, in principle, we could use formal methods to reveal insights in other normative domains, too.

As Hájek (MSc) notes, formal methods fall along a spectrum of formality. Representing an argument in *premise, premise,..., conclusion* form is on the low end of the formality spectrum. While it does schematize the argument in order to better reveal its logical structure, it doesn't significantly abstract from the content of the argument, and it's less a representation of the argument than a re-formulation of it. Probabilist modeling of degrees of belief with abstract logic or set-theory and algebra is on the higher end of the formality spectrum.

It's also important to note that formal modeling is distinct from some other uses of formal tools. For instance, formalism is sometimes used merely for the purpose of abbreviation, to avoid the labor of writing out lengthy principles and notions (Titelbaum 2013, p. 299). Philosophers often abbreviate claims and theses with logical symbolism. Consider, for example Williamson's (2000) thesis, which is often referred to as "E = K," that one's body of evidence is one's body of knowledge. Here Williamson isn't modeling, he's just employing logical symbols in a shorthand.

We also sometimes employ formalism not explicitly to model, but simply to aid in computation or the completion of some other task. For example, a would-be law student with basic training in symbolic logic might, while taking the LSAT, abbreviate sentences from a reasoning problem with sentence letters and Boolean connectives from a logical language in order to make the logical form of the problem clearer. In such a case, the test-taker is not modeling the logical structure of the problem in a very sophisticated or explicit way. Rather, the logical formalism is serving more like a shorthand or abbreviation to help the student reason more efficiently.

Similarly, we sometimes use formalism to perform other tasks by exploiting the similarity between the formal structure and some aspect of the world. For example, computer programmers regularly employ formal languages that are apt to model various features of the world. But often these programmers use their programming languages not explicitly to model, but to execute tasks based on the structural similarities of the language with features of the world.

### 4.1.2 The Aims of Formal Modeling in General

So, why bother with formal modeling? What kinds of insights can we gain by engaging in formal modeling? In this sub-section, I'll explain the purpose of formal modeling by way of surveying the benefits that formal modeling can bring. I note, however, that formal modeling is a double-edged sword. Below I present many of the virtues of formal modeling and the benefits it can bring, but there are corresponding vices and harms that can also arise from formal modeling. I mention some of these vices in passing, but I do not present them in detail.

One of the main benefits of formal modeling is that it promotes what Titelbaum (2013, p. 300) calls "good methodological hygiene."[123] The idea is that when formal modeling is done well, it requires us to isolate, simplify, and precisify our philosophical concepts and assumptions. This may be a virtue in its own right, but it also promotes critical reflection about our concepts and assumptions, and care in formally representing them. It requires us to think hard about how best to state our views and employ our concepts, and, in so doing, it can make perspicuous their flaws and virtues. Formal modeling can also reveal hidden assumptions and motivate us to augment our philosophical lexicon to fill in conceptual gaps.

---

[123]Several other philosophers have also noticed this phenomenon. See Hájek (MSc), Hansson (2000), and Wheeler (MS).

Formal modeling also promotes logical economy[124] by providing minimal and parsimonious formal representations of our concepts and assumptions in simple formal languages with minimal axioms. This encourages us to recognize the inter-reducibility of our philosophical concepts and assumptions. This in turn promotes good methodological hygiene and can help us gain better understanding of our philosophical notions and views, and their interrelations.

Employing formal models can also allow us to identify and gain insights about complex and delicate logical and mathematical relations among our concepts and claims by abstracting from their content. For instance, by representing arguments in a symbolic logic, we can recognize properties of our arguments without being concerned with whether or not the premises and conclusion are true. By appealing to a formal representation of an argument, one might better convince an interlocutor that a particular argument is poor because it is invalid even if the premises and conclusion are true.

In addition to making the logical and mathematical connections among concepts and claims easier to recognize, formal modeling can also make it easier to identify their properties and draw out their implications through inference. Indeed, as I'll argue below, a properly articulated formal modeling framework will include syntactical rules for generating new formulas within the formal system and modeling rules for generating insights about the modeled phenomena. In this way, formal modeling can give us a

[124]"Logical economy" is a modification of Hansson's "deductive economy" (2000, p. 165).

systematic, algorithmic means of drawing out the implications of our claims and gaining insights about the modeled phenomena. This can help safeguard us from error (Hájek, MSc).[125]

The logical economy and systematic, algorithmic nature of formal modeling also promotes a kind of completeness (Hansson 2000, p. 167). As I've noted above, it encourages us to determine all of the important features of the phenomena we're modeling, those features that turn out to be essential for the model to work properly. It also encourages us to draw out all of the important implications of our assumptions by making the process of doing so easier to accomplish.

Relatedly, formal representation can help us to identify previously unnoticed features of phenomena. In modeling some phenomenon, we might notice that the formal model is similar to the formal model of another phenomenon (Hájek MSc). Likewise, we might recognize common features of phenomena, and thereby recognize their common susceptibility to formal modeling, thereby gaining even further insights about the phenomena in question. So, we might thereby come to gain new insights about both phenomena through a form of theoretical cross-pollination.

Interestingly, we can also use formal modeling to guard against the potential vices of formalization. Done well, in the fashion I describe below, formal modeling, promotes the careful implementation of formal representations. It encourages us to be watchful for illicit applications of the formal representations to domains to which they don't

[125]Of course, if we are mistaken in generating these algorithms or inept in formally representing the phenomena, our "insights" will be of limited value.

apply. It also helps us be careful about drawing illicit conclusions about the modeled phenomena from the features of the models themselves. Furthermore, by employing multiple formal models of the same phenomena—especially by employing multiple formal models from different modeling frameworks—we can guard against oversimplifying our formal representations and reifying artifacts of the formal representations.

### 4.1.3 The Elements of a Formal Modeling Framework

Now that I have given a general account of formal modeling and characterized the benefits that it can bring, I want to characterize formal modeling in more detail by presenting the elements of well-crafted formal modeling framework. The account I describe here is adopted from Titelbaum (2013, Chs. 2-3).

First, following Titelbaum (2013, p. 11), I want to distinguish between specific formal models and the general formal modeling frameworks in which they are created. Probabilism is often characterized in conversation as a "formal model" of degrees of belief and degree of belief norms. But, as Titelbaum points out, this isn't quite right. Probabilism is more aptly described as a formal modeling framework, a "scaffolding" on which we can build many individual formal models, which represent particular cases, within the general formal modeling framework (2013, p. 11). For probabilism supports the creation of many particular models of the degrees of belief of many different agents at different times. Indeed, probabilism may be used to generate multiple models of a single individual's degrees of belief at a single time.

According to Titelbaum (2013, Ch. 2), a formal modeling framework has three main parts: a formal system, an interpretation of this system as it represents and applies to the modeled phenomenon, and a set of modeling rules which specify when and how the framework may be applied.[126] It's important to note that often philosophers and scientists do not present these elements of their formal modeling frameworks explicitly. Indeed, as Titelbaum notes, many philosophers do not think of probabilism and similar uses for formalism as instances of modeling. Nonetheless, when formal modeling is done well, it includes these three main components. And while it can be tedious to present all of the aspects of a formal modeling framework in detail (its formal system, for example), presenting these details is important for preventing abuse of the modeling framework.

The formal system consists of the formal parts of the framework that can be characterized without respect to their interpretation. In particular, the formal system includes a modeling language consisting of an alphabet of symbols and a set of formation rules that specifies which strings of symbols count as well-formed expressions of the language. The formal system also includes a set of what Titelbaum calls "systematic constraints," which apply to all models within the framework. In a probabilist modeling framework, the systematic constraints characterize the inferential

[126]As Titelbaum (2013, p. 17) acknowledges, the elements of a formal modeling framework could be carved up a little differently, but this account includes the key elements.

apparatus of the probability axioms and logical and mathematical transformation rules that allow us to derive consequences within the formal system.

A formal modeling framework's interpretation ties its formal system to the target phenomenon—the phenomenon it is intended to represent. An interpretation has three main parts: an account of the target phenomenon that is to be modeled, a scheme for representing this phenomenon within the formal system of the framework, and a scheme for applying insights derived in the formal system back onto the target phenomenon.

The main reason that we need to characterize the target phenomenon is to know what we're trying to model with our formal system. We need at least a basic characterization of the target phenomenon in order to get an interpretation up and running. We also need to characterize the target phenomenon in order to assess how well the formal system represents its features. Are all of the relevant features of the target phenomenon represented in the model? Is the representation skewed in important ways? Are there aspects of the formal model that do not correspond to any real features of the target phenomenon?

In many cases, the target phenomena of a formal modeling framework is empirically observable, and the modeling framework is designed to generate models that describe or make predictions about these phenomena. In many philosophical applications of formal modeling, like probabilism, things are a little more complicated. For philosophical modeling is often normative in addition to being descriptive. For example, decision theory is often construed as theory of how humans ought to behave given their degrees of belief and preferences. Additionally, in philosophical modeling, the purpose of modeling is not just to describe, make predictions, or generate verdicts about the target. Often these models are also intended to help characterize, explicate, or analyze some philosophical concept. For example, the target of a probabilist model is an agent's doxastic state at a given time. A probabilist model represents an agent's degrees of belief, and generates verdicts about whether these degrees of belief are coherent or rational. However the probabilist modeling framework is also supposed to shed light on the notions of *degree of belief*, *coherence*, and *rationality*—on what they are and what roles they play in an agent's life.

Apart from characterizing the phenomenon its models are about, a formal modeling framework's interpretation should also specify the connections between the formal system and the target phenomenon. Titelbaum advocates the importance of articulating what he calls the "representation scheme" and the "application scheme" of the framework. The representation scheme specifies how the target phenomenon is represented in the framework's models by elements of the formal system. The application scheme, on the other hand, specifies how we can gain insights about the target phenomenon by reading features of particular formal models back onto the target phenomenon. In the case of probabilism, the representation scheme tells us how an agent's degrees of belief are represented in models using a logical language and credence functions defined on this language. So, for example, one such principle might be:

| | |
|---|---|
| Representation of Certainty | If an agent, $\alpha$, is certain of a claim, $A$, then $cr_{\alpha}(A) = 1$. |

The application scheme tells us how to generate normative verdicts about an agent's degrees of belief based on our formal models. Here's one possible example of such a principle:

| | |
|---|---|
| Tautological Certainty | If $\vDash A$, then $\alpha$ ought to be certain of $A$. |

Titelbaum (2013, pp. 12-15) provides a clever triangular representation of how bridge principles connect a formal system, a target phenomenon, and a set of theoretical concepts.[127]

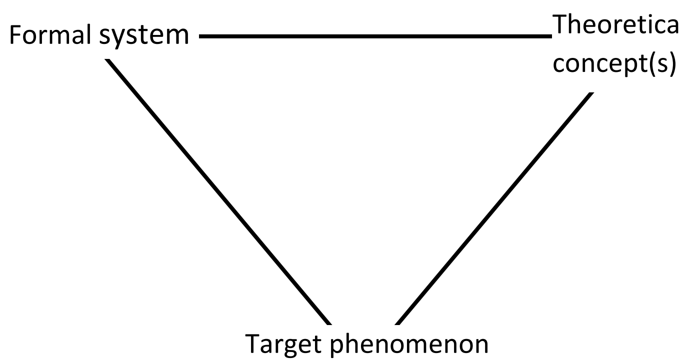


Figure 1: Formal Modeling Triangle

The edges represent bridge principles. A formal modeling framework's representation and application schemes connect the formal system and the target phenomenon, but we can also articulate bridge principles connecting the formal system and the target phenomenon directly to the theoretical concepts that the modeling framework is designed to represent and provide insights about.

[127]Titelbaum's (2013, p. 12) general formal modeling triangle connects a formal system, a philosophical concept, and a set of norms. I generalize the triangle to apply to descriptive/predictive modeling rather than to normative modeling only. I also make it explicit that a formal modeling framework could be designed to model a set of theoretical concepts, rather than just a single master concept.

Here's a version of the triangle for probabilism.[128]

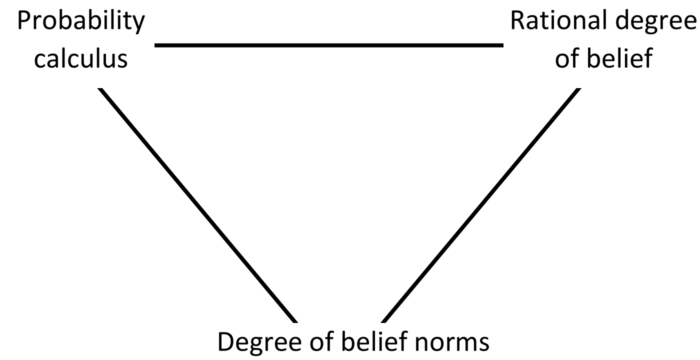


Figure 2: Probabilist Triangle

Probabilists use the formal system of the probability calculus to create models, which represent agents' degrees of belief in particular cases. Then they apply the formal models back onto the world by translating facts about the formal models into normative verdicts about the rationality of the degrees of belief represented in the models. This process helps us characterize general connections between the concept of *rational* or *coherent* degree of belief with the probability calculus and normative verdicts in particular cases.

In addition to the formal system and its interpretation, the third main part of a well-articulated formal modeling framework is a set of modeling rules, which specify when and how the framework should be applied to particular cases. These rules help

[128] This triangle is also a slight modification of Titelbaum's representation of probabilist triangle. It's worth noting that in his own Bayesian modeling framework, the Certainty-Loss Framework, Titelbaum attempts to connect a probabilistic formal system directly to a set of norms without an intervening philosophical concept (2013, p. 16). Titelbaum takes this approach not because he doesn't value insights about various philosophical concepts (like *rational degree of belief*), but because he wants to resist the view that all probabilist modeling should be aimed as providing an analysis of *probability* (or some other master concept) as is often assumed (2013, p. 16, n. 9).

specify what Titelbaum calls the framework's "domain of applicability," the set of phenomena that it can model adequately (2013, p. 85).[129] As Titelbaum notes, we shouldn't expect a single modeling framework to apply correctly to every instance of the target phenomenon (2013, p. 86). After all, modeling may require simplification and some attendant distortion. These distortions can cause the framework to yield systematically incorrect verdicts, descriptions, or predictions in some kinds of cases. So, to avoid being misled by the model, we need to specify modeling rules that demarcate the framework's proper domain of applicability. We also need rules to guide us in the proper use of the framework, especially in cases where different models built from the framework provide conflicting verdicts (Titelbaum 2013, p. 57, n. 3).

### 4.1.4 The Aims of Probabilist Modeling

Now that I've said more about what formal modeling is and why philosophers and scientists engage in it in general, I will revisit the aims of probabilist modeling in particular. I've already said quite a lot in the preceding discussion (especially Chapter 2) about why probabilists use credence functions to represent degrees of belief and the probability calculus to generate degree of belief norms. I won't rehash that material again here. Rather, my purpose in this sub-section is to characterize the aims of

[129] Fine (1973) uses the similar term, "domain of application," to refer to the set of phenomena for which the framework provides models. This usage is slightly different from Titelbaum's. For, as Titelbaum (2013, p. 85, n. 3) notes, there may be scenarios that a framework can represent in its models, but for which it will provide predictably incorrect verdicts (descriptions, predictions). Such scenarios are within the framework's domain of applica*tion* (in Fine's sense), but beyond its domain of applic*ability* (in Titelbaum's sense).

probabilist modeling so that I can provide a more unified account of probabilism's modeling methodology in accord with these aims, and so that I can appeal to these aims in my defense of particular methodological choices in the remainder of the chapter.

I think that probabilist modeling is oriented at three main aims: characterizing certain philosophical concepts including *degree of belief*, *coherent degree of belief*, and *rational degree of belief*; formulating general norms for degrees of belief; and generating verdicts about the coherence and rationality of degrees of belief in particular cases by applying these norms *via* probabilist models. Probabilists want to use their modeling framework to gain insights about the nature of degrees of belief and a normative standard of coherence for them. And they want to understand how this standard fits into broader theories of degree of belief rationality, and doxastic and practical rationality generally.

These aims are inter-related. Generating normative verdicts about agents' degrees of belief in particular cases requires the formulation of a set of norms that can be applied to particular cases. And the formulation of degree of belief norms depends on an understanding of the notion of *degree of belief* and a normative standard for degrees of belief. However, part of understanding the concepts *degree of belief*, *coherence*, and *rationality* is having a sense of what the norms for degrees of belief really are, and how these norms apply to particular cases.

Thus, I contend, probabilism's aims cannot be easily separated from one another. So, in addition to being designed to generate normative verdicts about particular cases, a probabilist modeling framework should also be designed explicitly to aid the characterization of *degree of belief*, *coherence*, and *rationality*, and the formulation of degree of belief norms. In the remainder of this chapter, I will begin to characterize such a probabilist modeling framework.

Before I move on, however, I want to take a moment to comment on the connection to the local approach to probabilist modeling mentioned in Chapter 3. On this approach, probabilists apply their formal system to produce many simple models of stylized descriptions of narrowly circumscribed cases in order to generate normative verdicts about those cases so described. The local approach is contrasted with the global approach, which is characterized as the attempt to produce a comprehensive, univocal model of the appropriate degrees of belief in every claim whatsoever, given all of the available evidence. Local modelers criticize the global approach based on the intractability of creating and employing such a model, and based on skepticism about the existence of a unique, correct model of the appropriate degrees of belief about everything.

Setting aside the issue of whether, in principle, there could be a univocal probabilist model in this sense, it's clear that the models we construct using the probabilist formalism must be local. For a comprehensive model of all of an agent's degrees of belief would be intractable. Even if we could construct such a model, it would be unwieldy; the costs of its use would not be worth the benefits. Thus, I admit, that probabilists must be local modelers in applying their formalism to generate verdicts about particular cases.[130]

However, I want to reiterate that the generation of verdicts about degrees of belief in particular cases is not the sole purpose of probabilism. So, probabilists can't be merely local modelers because, as I explain above, probabilist modeling frameworks must also be oriented towards characterizing concepts like *degree of belief*, *coherence*, and *rationality*, and formulating general degree of belief norms. In the remainder of this chapter, I provide a version of probabilism that aids in the characterization of these concepts and the formulation of degree of belief norms.

**4.2 CCF: A Probabilist Modeling Framework**

Now that I have provided a clearer account of the aims of probabilism and the elements of a well-articulated formal modeling framework, I will provide a probabilist modeling framework in accordance with these aims, with an eye towards avoiding or defanging the typical objections to probabilism presented in Chapter 3. I call this framework the Comparative Confidence Framework—CCF for short. I will characterize each of CCF's elements in accord with the account of these elements from the preceding sections. However, to make the formal system and its intended interpretation clearer, I won't present them separately. I will begin, in §4.2.1, by providing the characterization of degrees of belief that I will rely on in specifying the rest of the interpretation of CCF's

[130] Much as deductive logic is used to model deductive reasoning locally, rather than to represent every statement or proposition that a real agent may entertain, and all the deductive relationships among them.

formalism. (I favor this version for specific reasons, which I'll provide along the way.) Then, in §4.2.2, I will propose a scheme for representing degrees of belief within CCF. Finally, in §4.2.3, I will present a version of the qualitative probability calculus that, I will argue, generates appropriate degree of belief norms. I will leave my treatment of CCF's application scheme to Chapter 5.

### 4.2.1 Basic Characterizations of *Degree of Belief*, *Coherence*, and *Rationality*

Before I present CCF's formal system and its interpretation, I'll provide a basic characterization of the target phenomena it's meant to model. As I explained in the previous sub-section, probabilism is aimed at characterizing degrees of belief and a normative standard for them, formulating norms that characterize this standard, and applying these norms to generate verdicts about agents' degrees of belief in particular cases. So, in this sub-section, I will provide a basic characterization of *degree of belief* and its associated normative concepts of *coherence* and *rationality*, in order to inform my presentation of the other elements of CCF in the remainder of the dissertation.

I'll start with the notion of *degree of belief*. I won't attempt to analyze, explicate, or give an operational definition of this notion. As we saw in Chapter 2, each of the most prominent analyses, explications, and definitions faces significant difficulties. So, following Eriksson & Hájek (2007), I will endorse primitivism about *degree of belief* as a working assumption. That is, I will treat *degree of belief* as a primitive concept that cannot be faithfully analyzed or explicated in terms of other concepts. So, while I will not analyze, explicate, or define *degree of belief* in terms of anything more basic, I will provide a foundational characterization of degree of belief that will serve to clarify CCF's target phenomenon.

Without further ado, here is my approach. I characterize *degree of belief* as a gradational doxastic state. That is, degree of belief is a belief-like mental state about the way the world is. In other words, it is a mental state that has a "mind-to-world" direction of fit. In this sense, degree of belief is a representational rather than aspirational or conative state. As such, it is to be contrasted with other types of mental states, such as desire and intention.

Our degrees of belief are connected to—and (imperfectly) indicated by—our behavioral dispositions (including our betting behavior and tendencies to assert, exhibit surprise, etc.), our preferences, and other such mental states. But while our degrees of belief are connected to these other dispositions and states, I do not take them to be analyzable, explicable, or definable in terms of them. I take degrees of belief to be real, naturally occurring mental states in their own right.

Degree of belief is a *gradational* doxastic state in the sense that one's degree of belief about a claim can be weaker or stronger. One can be quite certain in a claim, pretty sure of it, unsure about it, pretty sure it is false, certain that it is false, etc. These differences in the strength of one's degrees of belief can be very fine-grained, and in some cases the strength of one's degree of belief can be identified very precisely and narrowly. In other cases, degrees of belief may be vague, fuzzy, or thick. So, the gradational structure of degrees of belief may be very fine-grained, but it may not be uniformly so.

Thus far, in characterizing degrees of belief, I have avoided commitment to a view about the objects of degrees of belief—e.g., as structured propositions, or sets of possible worlds. As we saw in Chapter 2, the issue is fraught with controversy, and tangled up with issues in the philosophy of mind and philosophy of language that would take us well beyond the scope of this dissertation. Though I won't commit to a view of precisely what the objects of degrees of belief are, I do want to characterize them well enough to model them formally. Mainly, I want to point out that whatever the doxastic objects are, they may be very fine-grained—fine-grained enough for an agent to have different doxastic attitudes about the truth of "Clark Kent is Superman" and "Clark Kent is Clark Kent," and about the truth of "this glass is full of water" and "this glass is full of H2O," and even perhaps to have different attitudes about "Neither Raquel nor Bernie got the job" and "Raquel didn't get the job and Bernie didn't get the job." So, in specifying CCF's formal system, we should choose a modeling language that can accommodate this fineness of grain. This fineness of grain is necessary if we are to represent (and model) the degrees of belief of non-ideal agents, who may be unaware of many metaphysical, logical, and, perhaps, even linguistic equivalences.

Characterizing degrees of belief and their objects is a large part of characterizing the target of probabilist modeling. But, as I explain at length above, we also need a basic understanding of the normative notion or notions that probabilist norms and verdicts are about. Once again, I won't attempt here to provide a complete account of all the normative notions that probabilism may address, but I do want to say enough to clarify the discussion to follow.

I'll distinguish two normative notions that are typically associated with probabilism: *coherence* and *rationality*. Probabilism, I take it, is intended primarily to characterize a notion of *coherence* among degrees of belief. This notion of *coherence* is akin to the notion of *consistency* among categorical beliefs. We may use the probability calculus to characterize *coherence* in much the same way that we use the propositional or predicate calculus to characterize logical consistency. This notion of coherence among degrees of belief is about how well an agent's degrees of belief regarding various statements (or propositions) fit together. The coherence of an agent's degrees of belief is distinct from the issue of how well-supported they are by the agent's evidence. Thus, assessments of the coherence of an agent's degrees of belief may contribute to assessments of the agent's rationality. However, such assessments do not always wholly determine the agent's rationality. So, in characterizing a notion of *coherence* for degrees of belief, we only partially characterize the notion of degree of belief *rationality*.

The notions of *coherence* and *rationality* are normative in the sense that they characterize what human agents' degrees of belief *ought* to be like. I will say more about the force of this "ought"—the sense of the normativity of these notions—in Chapter 5. The basic idea is that *coherence* and *rationality* characterize states that are epistemically good, right, obligatory, praiseworthy, etc. The norms that probabilism provides are norms about the ways in which our degrees of belief ought to hang together. And probabilist verdicts about the "rationality" of agents in particular cases are really verdicts about the coherence of the agents' degrees of belief in those cases.

#### 4.2.2 Representation Scheme

Now that I've given a basic account of the targets of probabilist modeling, I'll specify CCF's scheme for representing degrees of belief within its formal system. Along the way, I'll introduce CCF's modeling language. There are two key features of CCF's formal system and representation scheme that I will defend in this section. First, I will defend the use of a logical rather than set-theoretic modeling language on the basis that the logical approach better accommodates the representation of the degrees of belief of non-ideal agents. Second, I will defend a formal representation of degrees of belief that is fundamentally qualitative and comparative rather than quantitative.

##### 4.2.2.1 A Logical Representation of Doxastic Objects

The first component of CCF's formal system is a standard language for predicate logic (with identity, and perhaps functions), $\mathcal{L}$, containing countably many atomic sentences closed under the usual logical operations (as described in Chapter 2). In CCF's representation scheme, sentences of $\mathcal{L}$ are used in models to represent the objects of an agent's degrees of belief. Following my earlier convention, I will use $A, B, C, \ldots$as metalinguistic variables ranging over the sentences of $\mathcal{L}$. I choose to employ a language for predicate logic rather than a merely sentential language because it permits increased expressive power. However, I could have just as easily chosen a finite sentential language, which has the benefit of simplicity. However, for many applications, one might employ only a restricted part of the formal language—for example, the sentential part may suffice, and perhaps only a finite number of atomic sentences will be needed.

While CCF employs sentences of a logical language to represent doxastic objects, I want to reiterate that I am not assuming that sentences are the objects of degrees of belief. As I explained above, I will not take a stand in this dissertation about what the objects of degrees of belief really are. But, whatever they are, they are fine-grained enough that agents may have different doxastic attitudes about the truth of different natural language sentences that express the same proposition (Titelbaum 2013, p. 35). The sentences of a sentential language provide a fine-grained enough structure to allow the representation of agents who have different degrees of belief in metaphysically equivalent expressions, and in logically equivalent expressions as well.

The ability of logical approaches to represent the degrees of belief of agents who are ignorant of logical and metaphysical equivalences gives it a clear advantage over set-theoretic approaches in the context of using the probability calculus to model the degrees of belief of real agents. For, even if some set-theoretic approaches can represent agents who are ignorant of some equivalences between doxastic objects, every set-theoretic approach will fail to represent some cases of logical non-omniscience. This is because of the pure extensionality of sets: a set is nothing other than a collection of members. So, for example, consider the sets $A \cap B$ and $\overline{\overline{A} \cup \overline{B}}$. Even though these sets are expressed with different strings of symbols and formed by different operations, they are composed of the same members, and are, thus, identical. So, a set theoretic approach will be unable to represent an agent who is ignorant of this identity, and as a result may have distinct degrees of belief regarding them. Any set-theoretic representation of the objects of degrees of belief builds this kind of set-identity into its modeling language. So, any set-theoretic approach will be unable to faithfully model the doxastic objects of an agent for whom such identities do not hold.

It's worth noting, however, that some probabilists have suggested that this same feature—the pure extensionality of sets—makes set-theoretic approaches superior to logical approaches in the context of modeling logical omniscience. This is because, as Easwaran (2011a, pp. 323-324) explains, on a logical approach, the logical omniscience requirement will demand that an agent should be certain of infinitely many tautologies for a finite number of atomic sentences. Whereas, on a set-theoretic approach, the agent will be required to be certain of a merely finite number of doxastic objects. Thus, the set-theoretic approach appears to yield a less demanding version of the logical omniscience requirement. But, as Hawthorne (2009, p. 55, n. 8) points out, this appearance can be misleading. Even though the logical omniscience requirement on a (finite) set-theoretic approach will require certainty regarding only a finite number of doxastic objects, it builds a form of logical omniscience into the modeling language, since it builds set identities (and the recognition of these identities by agents) into the language.

So, while the extensionality of sets may appear to be an advantage in the context of demandingness concerns about the logical omniscience requirement, in fact, the set-theoretic approach faces a logical omniscience problem equally as difficult as that of the logical approach. But the set-theoretic approach has the added flaw that it is unable to faithfully represent the degrees of belief of logically non-omniscient agents. This is a major flaw, given the probabilist aim of using probabilist models to generate verdicts about the degrees of belief of real agents in particular cases. We can't use probabilist models to generate verdicts about agents (e.g. an agent's failure to recognize that two statements are logically equivalent) if we can't faithfully represent agents' degrees of belief in our models. So, even though the set-theoretic formulation of the probability calculus is somewhat simpler than the logical approach, in the context of the using the probability calculus to model non-ideal agents, the logical approach is superior. For this reason I will employ a logical modeling language in CCF's formal system.

#### 4.2.2.2 A Qualitative, Comparative Representation of Confidence

In addition to the logical language, $\mathcal{L}$, CCF's modeling language also includes a binary relation, $\succcurlyeq$, defined on the sentences of $\mathcal{L}$, which I call the "weak comparative confidence relation."[131] The weak comparative confidence relation is an order relation that compares pairs of sentences of $\mathcal{L}$. In CCF's representation scheme, ⌜ $A \succcurlyeq B$ ⌝ represents the modeled agent as at least as confident of $A$ as of $B$. In a particular model, all of the binary comparisons among sentences come together to form an ordering or ranking of sentences of $\mathcal{L}$. This ranking represents an agent's degree of belief set in the modeled scenario.

I further augment my modeling language with three additional comparative confidence relations defined in terms of $\succcurlyeq$: the strict comparative confidence relation, $\succ$, the comparative confidence equivalence relation, $\approx$, and the comparative confidence

[131] This aspect of my framework's representation scheme is adapted from Hawthorne (2009). Similar approaches are found in Savage (1972) and Joyce (1999). See Fine (1973), Fishburn (1986), Wong et al. (1991), and Capotorti & Vantaggi (2000) for overviews of comparative confidence approaches.

indeterminacy relation, $\sim$. ⌜ $A \succ B$ ⌝ represents that the agent as strictly more confident in $A$ than in $B$ ($A \succcurlyeq B$ but $B \not\succcurlyeq A$). ⌜ $A \approx B$ ⌝ represents that the agent is equally confident in $A$ and $B$ ($A \succcurlyeq B$ and $B \succcurlyeq A$). ⌜ $A \sim B$ ⌝ represents the agent's confidence that $A$ compared $B$ as indeterminate ($A \not\succcurlyeq B$ and $B \not\succcurlyeq A$). In addition to the comparative confidence relations, we also draw on a certainty predicate, $Cert(\cdot)$. ⌜ $Cert(A)$ ⌝ represents that the agent is certain that $A$.

So, in CCF, we model an agent's degrees of belief in a scenario by first assigning atomic sentences of $\mathcal{L}$ to represent the basic objects of the agent's degrees of belief. We then represent the agent's degrees of belief with a comparative confidence ordering defined on the fragment of $\mathcal{L}$ that consists of the set of atomic sentences assigned to the agent's basic doxastic objects closed under the usual logical operations. In particular, we represent the agent's certainties according to the scheme above. This ordering may include ties, and cycles, and it may be incomplete (some sentences of the modeling language might not be ranked).

Thus, CCF's representation scheme is fundamentally qualitative and comparative rather than quantitative. In CCF, a comparative confidence ordering plays the role of the numerical unconditional credence function in traditional quantitative probabilist modeling frameworks. This kind of qualitative approach has the advantage that it avoids the psychological implausibility and normative inappropriateness of the precise numerical representation of degrees of belief discussed in Chapter 3. For, on CCF's the qualitative approach, the fundamental representation of degrees of belief is merely a confidence ordering. This approach does not assume that a comparative confidence ordering will admit of any faithful representation with a numerical credence function—much less a precise, unique such function. For, as I note above, the ordering may include cycles that would preclude the faithful representation of an agent's degrees of belief with a function (or set of functions) that assigns a unique numerical value to each sentence. Furthermore, this representation is perfectly capable of representing imprecise or indeterminate degrees of belief for agents who simply don't have richly articulated comparisons of belief strength among statements. CCF is compatible with the imprecision and indeterminacy of the degrees of belief of real agents, due to their psychological limitations and, perhaps, their attempts to match the strength of their confidence with their imprecise, incomplete, or indeterminate evidence.

It's important to note, however, that just because the qualitative approach does not assume the numerical representability of degrees of belief, this does not mean that it denies the possibility of such representability. Rather, the qualitative approach helps to clarify the conditions under which numerical representations of degrees of belief are appropriate. If a confidence ordering has certain formal properties (including, for example, reflexivity, transitivity, and additivity), then it will be representable with a (set of) numerical credence function(s) that preserve its ordering: given $\succcurlyeq$, there will be an associated credence function, $cr(\cdot)$, such that, for all $A, B$, if $A \succcurlyeq B$, then $cr(A) \geq cr(B)$. If the ordering has certain additional properties, which I will discuss below, then the set of representing credence functions can be whittled down to a unique function that precisely represents the strength of an agent's confidence in each claim, and which captures the magnitude of the differences in strength of the agent's confidence among claims. Representing an agent's degrees of belief with a comparative confidence relation (at the most fundamental level of the representation) helps clarify when numerical representation is appropriate and which features of the numerical representation correspond to features of the modeled degrees of belief. Thus, the qualitative approach avoids excessive precision in its representation when it should do so, but it is also capable of representing numerically precise degrees of belief when agents have them.

In this connection, it's also important to note that the quantitative approach does not in fact have two advantages over the qualitative approach that it may at first seem to have.[132] The first apparent advantage of the quantitative approach is that it allows us to represent the magnitude of the differences in the strength of an agent's confidence among claims, whereas a qualitative approach can represent only the ordinal properties of the agent's confidence. The second apparent advantage of the quantitative approach is that it allows us to represent an agent's degree of belief in a particular claim in isolation, whereas the qualitative approach represents an agent's confidence in a particular claim only in comparison to other claims.

It's true that on the qualitative, comparative approach, an agent's degrees of belief are always represented in relation to other claims, not in isolation. And it's also true that a particular pairwise comparison, like $A \succ B$, merely represents that the agent is more confident in $A$ than $B$ without representing *how much* more confident the agent is that $A$ than that $B$. So, if the numerical approach can represent magnitudes of difference in

[132] See Levinstein (2013, Ch. 1) for discussion of these and other apparent advantages of the quantitative approach.

confidence, and if it can represent the strength of an agent's confidence in particular claims in isolation, then it does seem to have an advantage over the qualitative, comparative approach.

However, as I explain above, when a comparative confidence ordering representing an agent's degrees of belief has the right structural properties, it can be faithfully represented with a credence function—or set thereof—that may provide an apt representation of the magnitude of difference between an agent's confidence in different claims, and which may support representations of isolated degrees of belief in the form of expressions like $cr(A) = .9$, where the precision of this numerical value may be significant. So, a qualitative approach can represent magnitudes of the differences between an agent's degrees of belief in claims. When such a fundamentally qualitative approach can be legitimately supplemented with numerical credence functions that represent its confidence orderings, the representation of differences in relative strength between degrees of belief is made more perspicuous, and meaningful representations of agents' confidence in isolated claims becomes possible.

Furthermore, the quantitative approach is able to faithfully represent isolated degrees of belief and relative difference in strength among degrees of belief only when its credence functions induce comparative confidence orderings that have the kinds of properties mentioned above (reflexivity, transitivity, additivity, etc.). That is, the quantitative approach makes stronger assumptions than the qualitative approach, since each credence function provides a unique ordering among the degrees of belief of the agent, whereas a qualitative comparative confidence relation need not do so. Thus, as we will see below, assuming that real agents' degrees of belief can be faithfully represented with numerical credence functions—assuming that credence functions can represents agents' isolated degrees of belief and the relative strengths of their degrees of belief—is tantamount to assuming that real agents' degrees of belief can be mapped to real numbers between 0 and 1 in a way that makes them (already) probabilistically coherent. A qualitative approach need not do so, and so may represent cases where real agents depart from the ideal of numerical representability.

So, the quantitative approach has these apparent advantages only in the presence of some strong idealizing assumptions that are often inappropriate for the purposes of representing the degrees of belief of real agents. A fundamentally qualitative approach, on the other hand, promotes careful attention to the nuances of the degrees of belief of real agents. It also helps clarify when numerical representation of degrees of belief is appropriate, and helps to convey the significance of factors that contribute to the numerical representation in such cases. So, because a fundamentally qualitative approach avoids excessive numerical precision and clarifies when numerical precision is appropriate, that's the approach I will adopt below as I develop CCF.

### 4.2.3 Qualitative Comparative Confidence Axioms

Along with the logical modeling language and the comparative confidence relations, the third main element of CCF's formal system is a set of axioms for comparative confidence relations that plays an analogous role to that of Kolmogorov's axioms in traditional, quantitative probabilist modeling frameworks. Thus, these axioms specify a kind of coherence for comparative confidence relations that we can use to generate degree of belief norms via CCF's application scheme.

Just as there are several ways to axiomatize quantitative probability functions (several distinct sets of axioms that characterize the same probability functions), there are also several axiom systems for qualitative, comparative probability relations. However, whereas Kolmogorov's axioms for quantitative probability are accepted as standard, there is no single standard set of comparative probability axioms.[133] I will employ a set of axioms closely related to those of Hawthorne (2009).[134] I have chosen these axioms not so much because they are the simplest or most parsimonious, but because they are stated intuitively and in a way that makes it easy to draw connections to degree of belief norms. These axioms also clarify the connections between different properties of comparative confidence relations and the representability of these relations with quantitative probability functions. In the appendix, I present two alternative equivalent sets of comparative confidence axioms that are somewhat simpler, and preferable for some purposes. These alternative axioms also more closely resemble the most popular comparative probability axioms like those surveyed in Fine (1977), Fishburn (1986), Wong et al. (1991), and Capotorti & Vantaggi (2000).

Hawthorne specifies his axioms in two stages. First he states a set of seven axioms that specify what he calls "rudimentary confidence relations" (2009, p. 54). These

[133]For overviews of different axiomatizations of comparative probability, see, for example, Fine (1973), Fishburn (1986), Wong et al. (1991), and Capotorti & Vantaggi (2000).

[134] Hawthorne's axioms, in turn are based on those of Savage (1972).

axioms characterize a kind of quasi-order, a reflexive and transitive—but not necessarily complete—ordering of sentences of $\mathcal{L}$. Hawthorne explains that a rudimentary confidence ordering is induced by any probabilistic credence function, but he acknowledges that some rudimentary confidence functions may not be representable with probability functions (2009, pp. 58-59).

In the second stage, Hawthorne states an additional pair of axioms that place additional constraints on the rudimentary confidence relations. When a rudimentary confidence relation satisfies these additional axioms, it can be represented by a unique probabilistic credence function (Hawthorne 2009, p. 59-63). Then Hawthorne characterizes a notion of *proper extendability* of a rudimentary confidence relation to a relation that satisfies these additional axioms. He explains that the axioms for the properly extendable rudimentary confidence relations characterize an ordering structure that is induced by any probability function.[135] He also explains that every properly extendable rudimentary confidence relation can be represented by at least one probabilistic credence function—usually a set of such functions, all of which represent the confidence relation equally well. In what follows, I will provide a slightly modified, but equivalent version of Hawthorne's axioms.

We begin by characterizing the rudimentary confidence relations. I'll present the axioms and then provide a brief gloss of each. A relation, $\succcurlyeq$, defined on a logical language

[135] That is, for every probability function, $pr(\cdot)$, there is a properly extendable rudimentary confidence relation, $\succcurlyeq$, such that if $pr(A) > pr(B)$, then $A \succ B$, and if $pr(A) = pr(B)$, then $A \approx B$. I borrow the language of a function inducing the structure of a relation from Capotorti & Vantaggi (2000, p. 210).

$\mathcal{L}$ is a rudimentary confidence relation if and only if it satisfies the following axioms. For all $A, B, C, D \in \mathcal{L}$

0 $Cert(A)$ if and only if $A \succcurlyeq (A \vee \neg A)$ (certainty-confidence-connection);

1 $\neg(A \vee \neg A) \not\succcurlyeq (A \vee \neg A)$ (non-triviality);

2 $B \succcurlyeq \neg(A \vee \neg A)$ (minimality);

3 $A \succcurlyeq A$ (reflexivity);

4 If $A \succcurlyeq B$ and $B \succcurlyeq C$, then $A \succcurlyeq C$ (transitivity);

5.1 If $Cert(C \leftrightarrow D)$ and $A \succcurlyeq C$, then $A \succcurlyeq D$ (right equivalence);

5.2 If $Cert(C \leftrightarrow D)$ and $C \succcurlyeq B$, then $D \succcurlyeq B$ (left equivalence);

6.1 If for some $E$, $Cert\big(\neg(A \wedge E)\big)$, $Cert\big(\neg(B \wedge E)\big)$, and $(A \vee E) \succcurlyeq (B \vee E)$, then $A \succcurlyeq B$ (subtractivity);

6.2 If $A \succcurlyeq B$, then for all $G$ such that $Cert\big(\neg(A \wedge G)\big)$ and $Cert\big(\neg(B \wedge G)\big)$, $(A \vee G) \succcurlyeq (B \vee G)$ (additivity);

7 If $\vDash A$, then $Cert(A)$ (tautological certainty).

It's a little hard to explain the axioms without applying an interpretation to them. So, in what follows, I will describe how each axiom characterizes the abstract structural properties of the confidence relations, and I will follow up this abstract explanation with an intuitive interpretation of the axiom as it applies to degrees of belief. I want to stress, however, that these intuitive interpretations of the axioms are place-holders. I will put off the precise specification of CCF's application scheme, including its scheme for translating the comparative confidence axioms into degree of belief norms, until Chapter 5. There I will present various options for formulating degree of belief norms based on CCF's axioms, and I will show how choices among these options have a huge impact on the implications and plausibility of the resulting norms.

Hawthorne (2009, pp. 53-54) presents Axiom 0 as a definition, but perhaps it is better expressed as an axiomatic relationship between certainty and comparative confidence. It simply says that every sentence, $A$, that satisfies the certainty predicate is ranked at least as high as the tautology $(A \vee \neg A)$. Applied to confidence, this means that if one is certain that $A$, then one should be at least as confident that $A$ as that the simple tautology $(A \vee \neg A)$; and *vice versa*, if one is are at least as confident that $A$ as that $A \vee \neg A$, then one should be certain that $A$.

I'll present Axioms 1 & 2 together. Axiom 1, non-triviality, says that the simple tautologous sentence form $(A \vee \neg A)$ is never ranked lower than the simple contradictory sentence form $\neg(A \vee \neg A)$. That is, either $(A \vee \neg A)$ and $\neg(A \vee \neg A)$ are not compared in the ordering, or $(A \vee \neg A)$ is ranked strictly higher than $\neg(A \vee \neg A)$. However, Axiom 2, minimality, says that every sentence, $B$, is ranked at least as high as $\neg(A \vee \neg A)$, which holds the lowest place in the ordering. Substituting $(A \vee \neg A)$ for $B$, this means that $(A \vee \neg A)$ is ranked strictly higher than $\neg(A \vee \neg A)$ in the ordering. Thus, each rudimentary confidence ordering is non-trivial in the sense that at least one sentence, $(A \vee \neg A)$,[136] is ranked strictly higher than at least one other sentence, $\neg(A \vee \neg A)$. And since the minimality axiom invokes all sentences $B$ of $\mathcal{L}$, it guarantees that every sentence appears in the ranking—at least in comparison to simple contradictions of the form $\neg(A \vee \neg A)$, but not necessarily in comparison to any other sentences.

[136] One sentence *form*—of which there will be many instances—to be precise.

Applied to degrees of belief, Axioms 1 & 2 say that you should be strictly more confident in tautologies of the form $(A \vee \neg A)$ than you are in simple contradictions of the form $\neg(A \vee \neg A)$, and you should be at least as confident in any claim as you are in contradictions of the form $\neg(A \vee \neg A)$. The general spirit of these axioms, when applied to confidence, is that you should be more confident in simple tautologies than simple contradictions, and you should never be strictly less confident in any claim than you are in a simple contradiction.

Axiom 3, reflexivity, requires that each sentence is ranked at least as high as itself in the ordering. Given the definitions of $\approx$ and $\succ$, this entails that for each sentence $A$, $A \approx A$ and $A \nsucc A$. A sentence should never show up at more than one level in the ranking. Thus, reflexivity rules out certain kinds of cycles in the ordering. Like the minimality axioms, reflexivity also guarantees that every sentence $A$ of $\mathcal{L}$ appears in the ranking—compared, in this case, to itself but not necessarily any other sentence.

Applying the reflexivity axiom intuitively to confidence is a little tricky. One possible interpretation says that you should be at least as confident in any claim, $A$, as you actually are that $A$. From this it seems to follow that you should be exactly as confident as you are in every claim: if, for example, you *are* certain of $A$, then you *should be* certain of $A$. But this sounds like a kind of *a priori* vindication of each of your degrees of belief, and that can't quite be right. For there may often be many different degree of belief sets that all afford representation with confidence relations that satisfy the rudimentary confidence axioms. A better way to apply the reflexivity axiom to confidence is to interpret it as saying that, whatever your degree of belief in a particular claim is, you should be single-minded about it in comparison to other claims.[137] You shouldn't have a degree of belief structure that amounts to you being strictly more confident in a claim than itself.

Axiom 4, transitivity, says that if $A$ is ranked at least as high as $B$, and $B$ is ranked at least as high as $C$, then $A$ is ranked at least as high as $C$. Transitivity thereby extends the interdiction on cycles in the ordering put in place by the reflexivity axiom. Together, the reflexivity and transitivity axioms make the rudimentary confidence orderings what mathematicians call "quasi-orders" or "preorders" (Roberts 1985, p. 15; Hawthorne 2009, p. 56). The transitivity axiom translates into a pretty intuitive norm for confidence: if you're at least as confident in $A$ as you are in $B$ and you are at least as confident in $B$ as you are in $C$, then you should be at least as confident in $A$ as you are in $C$.[138]

Axiom 5.1 and 5.2, substitutivity of equivalences (left and right), says that when the sentence, $(C \leftrightarrow D)$, stating the material equivalence of $C$ and $D$, is ranked at least as high as $\neg\big((C \leftrightarrow D) \vee \neg(C \leftrightarrow D)\big)$, $C$ and $D$ should share the same place(s) in the ordering. Applied to confidence, Axioms 5.1 & 5.2 says that if you are certain in two claims have the same truth value, then your confidence judgments including them should agree. In particular, along with the reflexivity axiom, Axiom 5 says that if you are

[137] This does not mean that your degrees of belief should be sharp or precise.

[138] It's worth noting that, while transitivity has significant intuitive pull, it has been challenged. See Fishburn (1986, p. 339).

certain that two claims have the same truth value, then you should be equally confident in them.[139]

Axioms 6.1 & 6.2, subtractivity and additivity, collectively say that subtractions or additions of incompatible disjuncts don't make a difference in comparisons of pairs of sentences. According to subtractivity, if there is some $E$ for which $\neg(A \wedge E)$ is ranked at least as high as $\neg\big(\neg(A \wedge E) \vee \neg(A \wedge E)\big)$, $\neg(B \wedge E)$ is ranked at least as high as $\neg\big(\neg(B \wedge E) \vee \neg(B \wedge E)\big)$, and $(A \vee E)$ is ranked at least as high as $(B \vee E)$, then $A$ is ranked at least as high as $B$. Removing the incompatible disjunct, $E$, does not allow the ordering to reverse so that, say, $(A \vee E) \succ (B \vee E)$ but $B \succ A$. Similarly, according to additivity, if $A$ is ranked at least as high as $B$, then whenever there is some $G$ for which $\neg(A \wedge G)$ is ranked at least as high as $\neg\big(\neg(A \wedge G) \vee \neg(A \wedge G)\big)$, $\neg(B \wedge E)$ is ranked at least as high as $\neg\big(\neg(B \wedge G) \vee \neg(B \wedge G)\big)$, then$(A \vee G)$ is ranked at least as high as $(B \vee G)$. Adding the incompatible disjunct, $G$, doesn't change the order between $A$ and $B$. While the subtractivity and additivity axioms are a little more complex than the other axioms, they apply pretty straightforwardly to confidence: adding or subtracting disjuncts that you are certain are incompatible with claims in a confidence comparison should not change your confidence comparisons between the pairs of claims to which this incompatible disjuncts are added or from which they are subtracted.

[139] By reflexivity, $C \succcurlyeq C$ and $D \succcurlyeq D$. Assuming $Cert(C \leftrightarrow D)$, by Axiom 5.1 (or 5.2), $C \succcurlyeq D$ and $D \succcurlyeq C$. So, by the definition of $\approx$, $C \approx D$. Applied to confidence, this gives us the norm that if one is certain that two claims have the same truth value, then one should be equally confident in them.

Axiom 7, tautological certainty, places each logical truth, $A$, at least as high in the ordering as the simple tautologous form $\neg(A \vee \neg A)$. In conjunction with the other axioms, this places all logical truths at the top of the confidence ordering. As a confidence norm, tautological certainty says just what it sounds like: you should be certain of all logical truths. Thus, it is a logical omniscience requirement. Such requirements are controversial, as we saw in Chapter 3. And, in Chapter 5, I will discuss various ways that we can formulate less demanding versions of the tautological certainty norm. Ultimately, however, I will defend a version of the logical omniscience requirement. But I will put off that discussion for now.

Together, Hawthorne's axioms for the rudimentary confidence relations specify a kind of coherence very similar to Kolmogorov's probability axioms. As Hawthorne (2009, p. 55) notes, these axioms are actually a little bit weaker than typical sets of qualitative probability axioms, which usually include two additional kinds of axioms, a completeness axiom ($A \succcurlyeq B$ or $B \succcurlyeq A$, for all $A, B$) and an additional more complicated axiom the like of which I'll present below. Though the rudimentary confidence axioms are weaker than standard qualitative probability axioms, they characterize a structure that is induced by any Kolmogorovian quantitative probability function. That is, for any probability function, $pr(\cdot)$, there's a rudimentary confidence relation, $\succcurlyeq$, such that $A \succcurlyeq B$ just when $pr(A) \geq pr(B)$ (Hawthorne, 2009, p. 58). For a given $pr$ function, just define the relation $\succcurlyeq$ as follows: $A \succcurlyeq B$ if and only if $pr(A) \geq pr(B)$. Then check that, so defined, $\succcurlyeq$ satisfies the comparative confidence axioms.

While every probability function induces a rudimentary confidence relation, not all rudimentary confidence relations can be faithfully represented with probability functions (Hawthorne, 2009, p. 58). This is due in part to the fact that some rudimentary confidence relations may be incomplete: there may be some $A$ and $B$ for which $A \nsucc B$, $B \nsucc A$, and $A \not\approx B$. In many ways, this is a positive feature of the rudimentary confidence axioms: they provide coherence constraints for degrees of belief without making the strong assumption of completeness. For as Hawthorne (2009, p. 59) notes, the degrees of belief of many real agents may often be incomplete. This may be due to insufficient evidence, or cognitive or practical limitations. Indeed, in some cases, like Joyce's Black/Grey coins case discussed in Chapter 3, complete degrees of belief may be normatively inappropriate given the evidence. Or you may, for instance, simply have no good way of judging right now whether you should be more confident, less confident, or equally confident that it will rain tomorrow in Norman than that I currently have more than $5 in my wallet. So, the rudimentary confidence relations provide a handy tool for generating norms for incomplete degrees of belief.

However, Hawthorne (2009, pp. 59-60) points out that even if degree of belief completeness is too demanding in some cases, or otherwise normatively inappropriate, the notion of completeness may still have a role to play in characterizing norms for degrees of belief. For, as Hawthorne (2009, p. 59) points out, an incomplete confidence relation may satisfy Axioms 0-7, but there might be no way to coherently supplement the relation with additional comparisons to make it complete. In such a case, as Hawthorne puts it, the incompleteness of the relation hides an "implicit incoherence" in the relation (2009, p. 59). So, in the second stage of his axiomatization, Hawthorne characterizes a notion of *proper extendability* to preclude such cases of implicit incoherence.

Hawthorne characterizes the properly extendable rudimentary confidence relations as follows (2009, p. 60): a rudimentary confidence relation $\succcurlyeq_\alpha$ defined on a logical language $\mathcal{L}$ is properly extendable if and only if there is a rudimentary confidence relation $\succcurlyeq_\beta$ defined on $\mathcal{L}^+$ (an extension of $\mathcal{L}$) that agrees with the determinate part of $\succcurlyeq_\alpha$ ($\approx_\beta$ if and only if $\approx_\alpha$, and $\succ_\beta$ if and only if $\succ_\alpha$) on $\mathcal{L}$, and that also satisfies the following axiom for all sentences of $\mathcal{L}^+$:

(X) (i) (completeness) $A \succcurlyeq_\beta B$ or $B \succcurlyeq_\beta A$; and

(ii) (separating equi-plausible partitions)[140] if $A \succ_\beta B$, then for some integer $n$ there are $n$ sentences $S_1, \dots, S_n$ that $\beta$ takes to be mutually incompatible ($Cert_\beta\left(\neg(S_i \wedge S_j)\right)$ for $i \neq j$) and jointly exhaustive ($Cert_\beta(S_1 \vee \dots \vee S_n)$), and in all of which $\beta$ is equally confident ($S_i \approx_\beta S_j$ for each $i, j$) such that for each $S_k$, $A \succ_\beta (B \vee S_k)$.

As before, I'll explain Axiom (X) in terms of the abstract structure it requires, and then I'll provide an interpretation in terms of confidence. To satisfy Axiom (X), a confidence relation, $\succcurlyeq_\beta$, must be defined on a language enriched with a large equi-plausible partition. That is, the language must include a set of sentences, $\{S_1, \dots, S_n\}$, such that for each pair $i, j$ ($i \neq j$) of which $\neg(S_i \wedge S_j)$ is ranked at least as high as $\left(\neg(S_i \wedge S_j) \vee \neg\neg(S_i \wedge S_j)\right)$, $(S_1 \vee \dots \vee S_n)$ is ranked at least as high as $\left((S_1 \vee \dots \vee S_n) \vee\right.$

[140] A set $\{S_1, \dots, S_n\}$ of sentences is an $n$-ary equi-plausible partition for $\beta$ just in case $Cert_\beta\left(\neg(S_i \wedge S_j)\right)$ and $Cert_\beta(S_1 \vee \dots \vee S_n)$.

$\neg(S_1 \vee \ldots \vee S_n))$, and each sentence in the set has an equal place in the ranking. Axiom (X.i), completeness, says that every sentence in this enriched language must be compared to every other. So, there should be no gaps in the ordering. And Axiom (X.ii), separating equi-plausible partitions, says that whenever a sentence, $A$, is ranked strictly higher than another sentence $B$, the disjunction $(B \vee S_k)$ must be ranked strictly below $A$ for each $S_k$. If a rudimentary confidence relation satisfies Axiom (X), then whenever $A$ is ranked strictly above $B$, there is at least one disjunction $(B \vee S_k)$ that separates them in the ordering ($A \succ (B \vee S_k) \succ B$) (Hawthorne 2009, p. 60).

Applied to confidence, completeness says that there should be no holes in your degree of belief set. You should be opinionated—with some determinate degree of belief in every logical combination of the atomic claims toward which you have doxastic attitudes.

To explain how Axiom (X.ii), separating equi-plausible partitions, applies to confidence, Hawthorne suggests an example in which an agent, $\beta$, is more confident in $A$ than $B$, and there is a very large fair lottery with $n$ tickets. Additionally, Hawthorne assumes that $\beta$ is equally confident that each ticket will win, and certain that exactly one of the tickets will win (2009, pp. 60-61). In such a case, Hawthorne explains, Axiom (X.ii) requires that, since $\beta$ is more confident in $A$ than $B$, the lottery must have so many tickets that $\beta$ is more confident in $A$ than that either $B$ is true or any one particular ticket will win.

So, that's what it means for a rudimentary confidence relation to be properly extend*ed*, and for an agent's confidence to be representable with such a relation. A rudimentary confidence relation, $\succcurlyeq_{\alpha}$, is properly extend*able*, if, in addition to satisfying Axioms 0-7, there is a confidence relation, $\succcurlyeq_{\beta}$, defined on an a language enriched with equi-plausible partitions, which agrees with the determinate part of $\succcurlyeq_{\alpha}$ and which satisfies Axiom (X). So, the notion of degree of belief coherence Hawthorne characterizes holds that one's degrees of belief ought to conform to the norms suggested by Axioms 0-7. Additionally, while your degrees of belief need not be opinionated, and you may not have degrees of belief towards equi-plausible partitions like large fair lotteries, your degrees of belief should be supplementable with attitudes toward such doxastic objects such that they could be faithfully represented with a rudimentary confidence relation that satisfies Axiom (X).

Since every properly extendable confidence relation is a rudimentary confidence relation, it follows that every properly extendable confidence relation has a structure induced by any probability function. But, additionally, Hawthorne explains that every properly extendable confidence relation is representable with a set of probability functions such that for each one of them, $pr$, if $A \succcurlyeq B$, then $pr(A) \geq pr(B)$ (2009, p. 62). Additionally, Hawthorne points out that, if a confidence relation is properly extended, then there is a unique probability function $pr$ that represents it such that $A \succcurlyeq B$ if and only if $pr(A) \geq pr(B)$ (2009, p. 62).

Thus, Hawthorne's axioms provide a qualitative characterization of two kinds of coherence: a stronger notion associated with representation by properly extended confidence relations equivalent to that characterized by standard probabilism, and a weaker notion associated with representability by properly extendable confidence relations equivalent to that characterized by the extendability form of the probabilist norm discussed in Chapter 3. In Chapter 5, I will discuss which of these notions of coherence is most appropriate to the aims of probabilism. I will put off that discussion until then.

Now that I've presented the axioms in CCF's formal system, I want to present some of the benefits of qualitative axiomatizations in general and Hawthorne's axiomatization in particular. The first main advantage is that the qualitative approach is compatible with a fundamentally qualitative representation of degrees of belief. As I explain above, a fundamentally qualitative representation scheme avoids issues associated with the excessive precision of fundamentally numerical representations of confidence, clarifies the significance of numerical representations of confidence, and allows us to model non-ideal agents whose degrees of belief lack the structure necessary for faithful representation with a numerical credence function.

Similarly, qualitative axiomatizations clarify the comparative structure of the coherence requirement imposed by traditional quantitative probabilism. Hawthorne's approach, in particular, makes it clear what kind of comparative structure is necessary for faithful representation with a set of numerical probability functions or a unique probability function. Relatedly, by characterizing a weaker notion of coherence akin to the extendability form of the probabilist norm, Hawthorne's approach also has the advantage that it is compatible with the view that in some cases, like Joyce's Black/Grey Coins case, it is inappropriate to be opinionated or to have precise degrees of belief representable with point-valued credences. For, faithful representation with a properly extendable confidence relation is consistent with having incomplete or imprecise degrees of belief in the face of incomplete or imprecise evidence.

Finally, fundamentally qualitative approaches have the advantage that their axioms translate more easily into degree of belief norms. As we saw in Chapter 2, it can be difficult to specify the normative constraint that, for example, Kolmogorov's additivity axiom for numerical probability imposes. In general, as Capotorti & Vantaggi (2000) put it, most qualitative axioms permit "direct reading" unlike many numerical axioms. In particular, Hawthorne's axioms for confidence relations cry out for interpretation as confidence norms. Indeed, it was hard not to apply this interpretation to Hawthorne's axioms in glossing their meaning. This advantage is especially important to the probabilist aim of formulating clear norms for degrees of belief.

**4.3 Conclusion**

In this chapter, I provided an account of the nature and aims of formal modeling in general and probabilism in particular. I concluded that probabilism has three main aims: characterizing certain philosophical concepts (*degree of belief*, *coherence*, and *rationality*), formulating norms for degrees of belief, and generating normative verdicts about degrees of belief in particular cases. Thus, I argued that probabilism should be both local and global.

I also adopted and presented Titelbaum's (2013) account of the elements of a well-articulated modeling framework. And I set out to characterize the elements of a well-articulated probabilist modeling framework, CCF, beginning with what I call its methodological foundations. In particular, I set out a fundamentally qualitative, comparative version of probabilism built on a logical rather than set-theoretic modeling language.

I adopted a fundamentally qualitative, logical approach in CCF for several reasons. First, a fundamentally qualitative, logical representation scheme better allows us to represent the degrees of belief of real—non-ideal—agents within the modeling framework. This approach can model non-omniscience, and it avoids excessive precision and unrealistic assumptions. In general, the qualitative approach also helps clarify the significance of numerical representations of confidence via its representation theorems that specify connections between different properties of confidence relations and their numerical representability. From a normative perspective, this helps elucidate when it is appropriate to demand sharp degrees of belief based on the evidence. Furthermore, Hawthorne's qualitative approach in particular provides the advantage of flexibility in specifying different notions of coherence that make weaker and stronger demands appropriate to different circumstances.

In Chapter 5, I will lay out CCF's representation scheme, and thereby characterize its normative foundations. In particular, I will focus on the precise formulation of degree of belief norms based on Hawthorne's comparative confidence axioms. I will also briefly comment on the justification of these norms.

# 5 Normative Foundations

In Chapter 4, I began laying out a unified probabilist modeling framework, CCF (the Comparative Confidence Framework), which is intended to clarify probabilism's aims and implications, and to rebut or avoid the common objections to probabilism presented in Chapters 2 and 3. I provided an account of probabilism's aims, and I identified the key elements of a well-articulated probabilist modeling framework. Then I began to articulate CCF's details according to this account. I provided an account of the target phenomena that CCF is intended to model, namely degrees of belief, degree of belief coherence, and degree of belief rationality. I presented CCF's fundamentally qualitative (rather than quantitative) formal system consisting of a set of axioms for comparative confidence relations defined on a logical (rather than set-theoretic) modeling language. I also presented CCF's scheme for representing degrees of belief within CCF's formal system by means of comparative confidence relations. Finally, I sketched a scheme for translating CCF's comparative confidence axioms into norms for degrees of belief.

My purpose in this chapter is to continue the project of stating a unified probabilist modeling framework by providing an account of its normative foundations. In particular, my primary aim in this chapter is to state CCF's application scheme, its scheme for translating the comparative confidence axioms into degree of belief norms that can be applied to the degrees of belief of real human agents in particular cases. First, I will explain in more detail what a formal modeling framework's application scheme is, and why we need to specify one. Then I will lay out the core of CCF's application scheme, its scheme for translating the comparative confidence axioms and theorems into degree of belief norms. Once I have presented this scheme, I will present the resulting degree of belief norms together with the notion *of rational degree of belief coherence* that they encode. Finally, I will comment on the justification of CCF's norms in connection to the standard arguments for probabilism.

**5.1 What Is an Application Scheme and Why Do We Need It?**

As I explained in Chapter 4, a modeling framework's application scheme is the third component of its interpretation of its formal system. The first component is an account of the target phenomena to be modeled. The second component is the representation scheme, which specifies how to represent the target phenomenon within the formal system. The application scheme specifies how to project insights from within the formal system back onto the target phenomena. In a traditional probabilist modeling framework, the application scheme specifies how to translate the axioms and theorems of Kolmogorov's probability calculus into degree of belief norms. Probabilists then apply this scheme to models of particular cases in order to generate normative verdicts about agents' degrees of belief in those cases.

Why do we need an application scheme? Because, without one, the framework does not apply to the target phenomenon. On its own, without an interpretation, the formalism doesn't say anything about the degrees of belief of real agents. So, without an application scheme, probabilism doesn't generate degree of belief norms, and it doesn't provide verdicts about degrees of belief in particular cases.

As I noted in Chapter 4, probabilists don't usually state an application scheme explicitly. However, they rely on one implicitly. This implicit application scheme seems to hold that, by applying the representation scheme in reverse, we can translate the probability axioms and theorems directly into norms. Likewise we can translate derivations within models directly into normative verdicts about particular degrees of belief. This implicit application scheme imparts normative force to every axiom and theorem of the probability calculus, and it translates each derivation within a model into a normative verdict.

There are two main reasons why this implicit application scheme won't cut it. First, we can't read norms directly off of the axioms and theorems by substituting talk of comparative confidence relations (or about probabilistic credence functions) for talk about degrees of belief. This is because the axioms and theorems do not include any normative language. To turn the axioms into norms, we need to impart them with normative force. Thus, in the translation process, we need to introduce a normative operators like "should" or "ought." For example, a direct translation of the tautological certainty axiom yields:

- If $\vDash A$, then one is certain of $A$.

This is simply a descriptive claim.[141] Indeed (as applied to real agents) it is one that will very often be false. To turn it into a norm, we need to add a normative operator:

[141] That is, assuming "one" refers to a generic human agent. However, we could take "one" to refer to ideally rational agent. On this alternate reading, the translation of the axiom is normative in the sense that it is a description of an ideal agent, whom real

- If $\vDash A$, then one should be certain of $A$.

However, "should" and "ought" are not our only options for normative operators. For example, we can also translate the axioms into permissions.

- If $\vDash A$, then one is permitted to be certain of $A$.

Depending on which normative operator we choose to include in the translations, we generate very different norms with very different consequences. As I will show below, we also face many other options for generating norms from the comparative confidence axioms and theorems. Different choices among these options yield very different versions of probabilism, which have different features, and which are subject to different criticisms. We must specify CCF's application scheme explicitly in order to make its consequences clear. As Titelbaum (2013, Chs. 4-5) points out, we can avoid some of the typical objections to probabilism just by stating our application scheme carefully.

Second, we also need to state CCF's application scheme explicitly in order to avoid imparting normative force to skewed aspects of the formal system. As I explained in Chapter 4, formal modeling is fruitful in large part because it allows us to abstract from the content of the target phenomena to view their formal features. The process of formalization typically includes some simplification and regimentation of the target phenomena, which can introduce some bias into the formal representation. No formal

human agents *ought* to emulate. This kind of blanket approach to normativity doesn't permit the kind of nuance that the following treatment provides.

representation is perfect, after all. Because of this, there may be aspects of our formal model that we don't want to translate back onto degrees of belief with normative force.

We've already modified the typical probabilist modeling framework to avoid imparting illicit normative force to its numerical representation of degrees of belief. However, even so, there may be features of CCF's formal system that we don't want to impose onto degrees of belief as norms. For example, for the reasons presented in Chapter 3, we might want to avoid imposing the kind of logical omniscience requirement suggested by the tautological certainty axiom. So, instead of translating this axiom as

- if $\vDash A$, then one should be certain of $A$

we might want to translate it as

- if $\vDash A$, then, if one has any degree of belief in $A$, one should be certain of $A$.

Instead of demanding full logical omniscience, this second translation of the norm makes the weaker demand of logical infallibility.

So, we need to specify CCF's application scheme in order to make its normative consequences clear and to avoid imparting normative force to aspects of the formalism that are mere artifacts of that formalism. In the next two sections, I will present the array of options we face in formulating probabilist norms from CCF's axioms and theorems, and I will evaluate these options in light of probabilism's aims.

## 5.2 Norm Formulation Options

Recently, several philosophers have turned their attention to the question of precisely how we should formulate coherence norms based on formal systems like deductive logic or the probability calculus.[142] The contributors to this literature have identified four main parameters that can be varied in norm formulation.[143]

1. **Normative operator type:** What type of normative operator does the norm include? Is it a deontic operator: obligation, permission, or interdiction? Or does it specify a *pro tanto*, defeasible recommendation of reason? Is the normative operator strict—like obligation—or slack—like the reason operator?

2. **Normative operator scope:** When the norm is logically complex, what is the scope of the normative operator? Are there multiple normative operators within the norm?

3. **Normative operator polarity:** Is the polarity of the normative operation positive or negative? Does the norm tell one to *adopt* a particular doxastic attitude (positive) or to *avoid* adopting one (negative)?

4. **Knowledge condition:** Is the norm conditional on the agent's knowledge of some fact—logical facts, in particular?

I will make two modifications to this set of parameters. First, I'll add an additional operator to the list of possible normative operators. MacFarlane, Broome, *et al*., focus on the reason operator and the deontic operators. Since many probabilists hold that the probabilist norm is an evaluative norm, I'll add an evaluative operator to the list, namely the goodness operator, "it is good that."[144] Second, I'll add an additional parameter to

[142] The roots of this growing literature can be found in, for example, Harman (1986), Broome (1999), MacFarlane (MS), and Field (2009).

[143] This list of parameters is due to MacFarlane (MS). MacFarlane's presentation of the formulation options is the most complete and systematic. His paper concerns the formulation of norms for categorical belief based on deductive logic. In what follows, I will draw heavily on MacFarlane's work, adapting his approach to the task of generating degree of belief norms from the comparative confidence axioms.

[144] I could also add additional evaluative operators, such as a badness operator, an indifference operator, or comparative operators – e.g., a better-than and a worse-than operator. For the sake of simplicity, I'll focus on the goodness operator here.

the list, which I call the "attitude condition." This is the condition present in the modified version of the tautological certainty norm given above:

- If $\vDash A$, then, if one has any degree of belief in $A$, one should be certain of $A$.

When this condition is present, the norm applies on the condition that one has an attitude (a degree of belief or confidence comparison) toward the claim(s) in question. In the case of the modified version of the tautological certainty norm, the addition of the attitude condition changes the norm from a requirement of logical omniscience to a requirement of logical infallibility with regard to the claims towards which one has an attitude.

By varying these parameters, we can generate a wide array of formulations for any given norm—as many as 96 different formulations for some norms. So, before I explain the parameters in more detail, and present the different options we can generate by varying them, I want to introduce a scheme for naming the formulations. I hope that this naming scheme will facilitate and clarify the discussion to follow.[145]

According to this scheme for naming the formulations, each formulation is named by a string of initials that specify its scope, normative operator, polarity, and whether it is conditional on logical knowledge or possession of some attitude. The first initial of each name is a capital letter, `C`, `B`, or `W`, to indicate the scope of its normative operator(s). I assume here that all logically complex statements of norms employ (material) conditionals as their main connectives. So, the statement of a norm can

[145] I modify MacFarlane's scheme (MS, p. 7).

include a normative operator in the consequent alone (C), in both antecedent and the consequent (B), or the operator can have wide scope over the whole conditional (W). The second initial of each name is a lower case letter, o, p, r, or g, to indicate which type of normative operator the norm includes, obligation (o), permission (p), reason (r), or goodness (g).[146] The third initial of each name is either a plus sign (+) or a minus sign (−), to indicate that the normative operation has positive or negative polarity.

Every norm's name includes at least these first three initials, to indicate its operator(s), scope, and polarity. When a norm is not logically complex, its name will always indicate that its operator has widest scope (W). In addition to the first three initials, a formulation's name may also include one or two suffixes: a k to indicate that the norm is conditional on knowledge of some logical fact, or an a to indicate that the norm includes an attitude condition. With this naming scheme in hand, I will now present the parameters and the norms that they allow us to generate in more detail.

**5.2.1 Operator Type**

The first parameter I'll consider is the type of normative operator the norm includes. We have a choice among four main options: obligation, permission, reason, and goodness. I don't include interdiction as an option in its own right since we can conveniently state interdictions in terms of obligation or permission. There are a few main differences that distinguish the operators. First, some of them are strict, while

[146] I assume here that a single norm will not include multiple normative operators of different types. So, for example, I will ignore formulations like the following: if it is good for one to be at least as confident in $A$ as in $B$, and at least as confident in $B$ as in $C$, then one is obligated to be at least as confident in $A$ as in $C$.

others are slack. A normative operation is strict if one is subject to legitimate criticism when the norm says to φ, but one fails to φ. A normative operation is slack if one is not necessarily subject to criticism when the norm says to φ, but one fails to φ.

The deontic operators, obligation and permission, are strict. Recommendations of reason, on the other hand, are slack. For example, consider the `Co+`, `Cp+`, and `Cr+` forms of the tautological certainty norm.

`Co+` If $\vDash A$, then one is obligated to be certain of $A$.

`Cp+` If $\vDash A$, then one is permitted to be certain of $A$.

`Cr+` If $\vDash A$, then one has reason to be certain of $A$.

If one is uncertain of a particular tautology, $A$, one violates the `Co+` form of the norm, but not necessarily the `Cr+` formulation. This is because recommendations of reason are defeasible and *pro tanto*.[147] So, one's uncertainty in a tautology may be consistent with satisfaction of the `Cr+` formulation if there are other reasons that defeat or over-ride one's reason to be certain. For example, the tautology in question might be too complex for one to recognize it as such, or one might possess misleading evidence that leads one to think the tautology is false. As I noted above, permissions are also strict. The `Cp+` form of the norm permits certainty in any tautology regardless of one's evidence or the complexity of the tautology.

[147] When I say that reasons are *pro tanto*, I mean that a given reason might not be conclusive. When I say that reasons are defeasible, I mean that additional considerations may remove or mitigate the force of the reason.

It's not as clear whether the goodness operator is strict or slack. Consider the `Cg+` form of the tautological certainty norm.

`Cg+` If $\vDash A$, then it is good for one to be certain of $A$.

There's a sense in which it is strict: if one is uncertain of a tautology, then one's degree of belief could have been better. So, if one violates the norm, then one merits a negative evaluation. However, there's also a sense in which the goodness operator is slack. For, there seem to be cases where one's uncertain degree of belief in a tautology could be good. I have in mind again cases in which the tautology is extremely complex and the agent possesses misleading evidence that suggests that the tautology is false. One might argue that, in such a case, it is not only good, but best for the agent to be uncertain of the tautology.

I think that the goodness operator is strict, but that it gives the appearance of being slack in some cases. This is because we can make evaluations of goodness from different perspectives. We can make evaluations from an all-things-considered perspective, or we can evaluate based on subsets of our values. Evaluations of goodness, I contend, are strict relative to the set of values they are based on. Thus, it is important to keep in mind the set of values on the basis of which one is conducting an evaluation. From the perspective of coherence, it is perhaps good for one to be certain of all tautologies. But from a perspective that also takes evidential considerations into account, it might not always be good for one to be certain of a tautology, depending on one's evidence.

In addition to strictness, another feature that distinguishes the operators is how they are constrained by our limitations—to what extent they are subject to "ought"-implies-

"can" constraints. It is widely accepted that obligations are constrained by our abilities: one cannot be obligated to do what one cannot do. It is also widely accepted that evaluative norms are not subject to "ought"-implies-"can" constraints. There's a sense in which it would be good for me to be logically omniscient, though I cannot be. If there were a logical omniscience pill (with no major side effects), I would be a fool not to take it. Recommendations of reason are not directly constrained by our abilities either. Rather, our limitations provide competing reasons. So, in the scenario in which I am considering an extremely complex tautology, I have reason to be certain of the tautology per the `Cr+` form of the tautological certainty norm, but my cognitive limitations also provide a countervailing reason to be uncertain.

How do the different operations relate to each other? Holding the other parameters fixed, obligation formulations entail permission and reason formulations. If one is obligated to $\varphi$, it must also be permissible for one to $\varphi$. Likewise one must also have some reason to $\varphi$. Plausibly, obligation formulations also entail goodness formulations. If one is obligated to $\varphi$, there is at least some perspective—perhaps the all-things-considered perspective—from which it is good for one to $\varphi$.

Other things equal, permissions do not seem to entail or be entailed by recommendations of reason. A permission to $\varphi$ doesn't necessarily give one any reason to $\varphi$, or indicate that it is good for one to $\varphi$. Likewise, if one has reason to $\varphi$, it's not necessarily permissible for one to $\varphi$. There might be countervailing reasons that make it impermissible for one to $\varphi$.

The relations between recommendations of reason and goodness norms are not as clear. If one has reason to $\varphi$, there is at least some sense in which it is good for one to $\varphi$. But it doesn't follow that, if one has reason to $\varphi$, then it is good for one to $\varphi$, all things considered. Similarly, if there's a sense in which it's good for one to $\varphi$, then there must be at least some reason for one to $\varphi$, *pro tanto* and defeasible though it may be. Despite these apparent entailment relations, it's not clear that reasons formulations and goodness formulations are equivalent. After all, the reasons operator is slack, but the goodness operator is strict relative to a set of values.

### 5.2.2 Scope

When our norms are logically complex, we have three choices with regard to the scope of their normative operators: the normative operator can be embedded in the consequent of the conditional, as in the `Co+` form of the transitivity norm; there can be normative operators in the antecedent and the consequent, as in the `Bo+` form; or the normative operator can have wide scope over the whole conditional as in the `Wo+` form.

`Co+` If one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is obligated to be at least as confident in $A$ as in $C$.

`Bo+` If one is obligated to be at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is obligated to be at least as confident in $A$ as in $C$.

`Wo+` One is obligated to see to it that if one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is at least as confident in $A$ as in $C$.[148]

[148] Following MacFarlane (MS), I use the phrase "see to it that" in the wide scope formulations to clarify the scope of the normative operator. Compare "one is obligated

The scope of the normative operator makes a huge difference to the implications of a norm. In the Co+ form of the norm, the normative operator has narrow scope over the consequent of the conditional. So, when the antecedent is satisfied, the normative conclusion in the consequent "detaches" (Broome, 1999, p. 401). Thus, when one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, one becomes obligated to be at least as confident in $A$ as one is in $C$, regardless of whether one's confidence comparisons of $A$ and $B$ and $B$ an $C$ are well-founded.

Wide scope norms like the Wo+ form of the transitivity norm, on the other hand, don't generate determinate normative conclusions like narrow scope norms. According to the Wo+ form of the transitivity norm, if one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is obligated either to be at least as confident in $A$ as in $C$, or one is obligated to revise one's comparisons of $A$ and $B$ and $B$ and $C$. Wide scope norms don't tell one whether to embrace the consequences of one's degrees of belief or revise one's degrees of belief.

Narrow scope norms with normative operators in both the antecedent and the consequent, like the Bo+ transitivity norm, make still a different kind of demand. They extend normative pressure one is already under (MacFarlane MS, p. 10).

Bo+ If one is obligated to be at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is obligated to be at least as confident in $A$ as in $C$.

---

to see to it that, if $A$, then $B$" with the more awkward formulation, "one is obligated that, if A, then B." By using "see to it that" in this way, I do not mean to suggest that we have direct voluntary control over our doxastic states.

The `Bo+` transitivity norm says that, if one is already obligated to be at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is also obligated to be at least as confident in $A$ as in $C$. The norm doesn't apply any normative pressure to one, if one not obligated to be at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$.

### 5.2.3 Polarity

The polarity of the normative operation in a norm determines whether the norm requires (or recommends) us to *adopt* some attitude or *avoid* adopting one. Compare the `Wr+` and `Wr−` forms of the entailment norm.[149]

`Wr+` If $A \vDash B$, then one has reason to be at least as confident in $B$ as one is in $A$.

`Wr−` If $A \vDash B$, then one has reason not to be more confident in $A$ than $B$.

The `Wr+` form has positive polarity. If $A \vDash B$, it recommends that one has reason to be at least as confident in $B$ as one is in $A$, regardless of whether or not one currently has attitudes toward $A$ and $B$. The negative polarity form, `Wr−`, on the other hand, recommends that, if $A \vDash B$, one has reason not to be more confident in $A$ than $B$. This is perfectly consistent with one having no attitudes toward $A$ and $B$ whatsoever. Thus, the `Wr−` form of the norm does not recommend that one adopt new attitudes; it merely

[149] Note that even though the `Wr+` and `Wr−` forms of the entailment norm take the form of conditionals, and the normative operators appear only in the consequent, I still call them wide scope formulations. This is because even though the norms take the form of conditionals, I think of them as norms conditional on logical facts. In each case, the norm in the consequent of the conditional is not logically complex, so the normative operator trivially has widest scope in each case. For an example of a `C` norm conditional on a logical fact, consider the `Cr+` formulation of the entailment norm for categorical belief: If $A \vDash B$, then if one believes $A$, then one has reason to believe $B$.

recommends coherence among the attitudes one already has. In this sense, the negative polarity form is less demanding than the positive polarity form.

In many cases, adopting the negative polarity form of a norm is a good way to make the norm less demanding. In this way, it captures part of the idea behind the extendability form of the probabilist norm discussed in Chapter 3. However, in some cases, the positive and negative polarity forms of a norm are equivalent. Compare the `Wr+` and `Wr−` forms of the tautological uncertainty norm.

> `Wr+` If $\vDash A$, then one has reason to be certain of $A$.
>
> `Wr−` If $\vDash A$, then one has reason not to be uncertain of $A$.

A recommendation to be certain of all tautologies is equivalent to a recommendation not to be uncertain of any tautology. In cases such as this, we must adopt some other formulation in order to weaken the demands of the norm.

Holding other parameters fixed, positive polarity forms entail their negative polarity counterparts. For example, compare the `Wo+` and `Wo−` forms of the entailment norm.

> `Wo+` If $A \vDash B$, then one is obligated to be at least as confident in $B$ as one is in $A$.
>
> `Wo−` If $A \vDash B$, then one is obligated not to be more confident in $A$ than one is in $B$.

If one is obligated to be at least as confident in $B$ as one is in $A$, then one is obligated not to be more confident in $A$ than one is in $B$.

In general, the negative polarity forms do not entail their positive polarity counterparts. For example, comparing the `Wo+` and `Wo−` forms of the entailment norm again, an obligation not to be more confident in $A$ than $B$ doesn't necessarily mean that one should be at least as confident in $B$ as in $A$. It could also mean that one is permitted to have no confidence comparison between $A$ and $B$. However, as we saw above, in some cases, like the tautological certainty norm, the positive and negative polarity forms are equivalent.

Wr+ If $\vDash A$, then one has reason to be certain of $A$.

Wr− If $\vDash A$, then one has reason not to be uncertain of $A$.

Any positive polarity norm to be certain in a claim will be equivalent to the negative polarity norm not to be uncertain in that claim.

**5.2.4 Knowledge Conditions**

So far I've presented options for manipulating the normative operators in our norms—their type, scope, and polarity. However, in some cases we can manipulate additional aspects of the norm. Some norms, for example the tautological certainty norm and the entailment norm, are conditional on logical facts. Consider the Wo+ forms of these norms.

| | |
|---|---|
| Tautological Certainty (Wo+) | If $\vDash A$, then one is obligated to be certain of $A$. |
| Entailment (Wo+) | If $A \vDash B$, then one is obligated to be at least as confident in $B$ as one is in $A$. |

These norms say, respectively, that one is obligated to be certain of all tautologies, and one is obligated to be at least as confident in any claim as one is in the claims that entail it.

These norms are very demanding. They fall far afoul of "ought"-implies-"can" constraints. To make them less demanding, rather than making them conditional upon bare logical facts, we can make them conditional upon *knowledge* of these facts, as in the Wo+k forms.

Wo+k If one knows that $\vDash A$, then one is obligated to be certain of $A$.

Wo+k If one knows that $A \vDash B$, then one is obligated to be at least as confident in $B$ as one is in $A$.

These norms make much weaker demands. The Wo+k form of the tautological certainty norm demands certainty in the tautologies one knows as such. And the Wo+k form of the entailment norm demands that one is at least as confident in entailed claims as one is in the claims that one knows entail them. Adding knowledge conditions to norms of this kind can bring them more nearly within reach of the abilities of real agents.

Since knowledge is factive, when we hold other parameters fixed, formulations without knowledge conditions entail their counterparts with knowledge conditions, but not *vice versa*. For example:

Wg+ If $A \vDash B$, then it is good for one to be at least as confident in $B$ as one is in $A$.

Wg+k If one knows that $A \vDash B$, then it is good for one to be at least as confident in $B$ as one is in $A$.

Given the Wg+ for of the entailment norm, it follows that Wg+k form must hold, but not *vice versa*.

**5.2.5 Attitude Conditions**

An alternative way to make norms less demanding—even when they aren't conditional upon logical facts—is to include an attitude condition. Compare, for example, the Wo+ and Wo+a forms of the tautological certainty norm.

Wo+ If $\vDash A$, then one is obligated to be certain of $A$.

Wo+a If $\vDash A$, then, if one has a degree of belief in $A$, one is obligated to be certain of it.

Rather than requiring logical omniscience, like the Wo+ form, the Wo+a form requires mere logical infallibility: one might not be certain of all tautologies, but one is certain of all tautologies one invests confidence in.

Thus, attitude conditional formulations of the norms capture the spirit of the extendability form of the probabilist norm mentioned in Chapter 3. Namely, probabilism doesn't require one to be opinionated, but it does require one's degrees of belief to be coherently augmentable with additional degrees of belief. Thus, attitude conditional formulations are similar to negative polarity formulations except that in the case of certainty norms, as we saw above, negative polarity formulations are equivalent to their positive polarity cousins. In these cases, attitude conditions can be added to achieve the desired diminution of demandingness. However, as I will explain in the next sub-section, there are also cases where negative polarity formulations are more natural than attitude conditional formulations for this purpose.

When we hold other parameters fixed, formulations that do not include attitude conditions entail their counterparts that do, but not *vice versa*. For example, consider the Wo+ and Wo+a forms of the tautological certainty norm.

Wo+ If $\vDash A$, one is obligated to be certain of $A$.

Wo+a If $\vDash A$, one is obligated to see to it that, if one has a degree of belief in $A$, then one is certain of it.

In most cases, when other parameters are held fixed, the negative polarity form of a norm is equivalent to the positive polarity form with an attitude condition. So, in these cases, the negative polarity form and the attitude condition make each other redundant. Certainty norms are the exception. In those cases, as we have seen, the negative and positive polarity forms of the norm are equivalent, other things equal. In these cases, the addition of an attitude condition makes the norm less demanding, as we have seen.

In formulations that include knowledge conditions, these conditions seem to make the inclusion of attitude conditions redundant. For example:

`Wo+k` If one knows that $\vDash A$, one is obligated to be certain of $A$.

`Wo+ka` If one knows that $\vDash A$, one is obligated to see to it that, if one has a degree of belief in $A$, then one is certain of it.

It is widely—if not universally—accepted that knowledge in a claim entails belief in that claim. So, if one knows that $A$ is a tautology, it seems to follow that one believes $A$ is a tautology. Setting aside the precise connection between categorical belief and degree of belief, this suggests that, if one knows that $A$ is a tautology, then one has some degree of belief in $A$. Thus, the inclusion of the knowledge condition in the norm seems to make the attitude condition redundant.

However, the connections between knowledge and other mental states are controversial. Consider the `Wo+k` form of the tautological certainty norm on its own.

`Wo+k` If one knows that $\vDash A$, one is obligated to be certain of $A$.

On some understandings of the connection between degree of belief, categorical belief and knowledge, this form of the norm is trivial.[150] For, some epistemologists hold that knowledge of a claim entails certainty in that claim.[151] If this is the case, then, if one knows $A$ is a tautology, then one is certain that it is. It would be strange for one, then, to be certain that $A$ is a tautology, but uncertain in $A$. As I will argue below, because of difficulties associated with the analysis of knowledge, it may be best to avoid knowledge conditional formulations altogether.

### 5.3 Evaluation of the Formulation Options

Given the options, how should we formulate our probabilist norms? In this section, I will weigh the advantages and disadvantages of each option. In the end, I will argue for a hierarchy of wide-scope, evaluative, formulations. First, I'll argue that the formulations should be wide(st) scope. Then I will argue for formulations with the evaluative goodness operator. Finally, I will argue that probabilists should want to characterize higher standards of coherence as well as lower standards. For this reason, I propose a hierarchy of evaluative formulations.

#### 5.3.1 Scope Issues

Following MacFarlane (MS, pp. 9-10), I'll start by ruling out the narrow scope formulations. We can rule out these formulations on the basis that, with the exception

---

[150] Thanks to Martin Montminy for pointing this out to me.

[151] Indeed, some epistemologists have suggested that the connection between categorical belief and degree of belief is that categorical belief in a claim is certainty (maximal degree of belief) that the claim is true. See Christensen (2004, Ch. 2) for discussion.

of attitudes toward logical truths and falsehoods, logic alone doesn't recommend or require adopting or avoiding particular attitudes. Rather, it tells you to maintain coherence among your attitudes. Consider the Co+ formulation of the transitivity norm.

> Co+ If one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is obligated to be at least as confident in $A$ as in $C$.

If one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then the Co+ form says that one is obligated to be at least as confident in $A$ as one is in $C$, regardless of whether or not one's comparisons of $A$ and $B$ and $B$ and $C$ are well-founded. In some cases, as I explain above, one should revise these comparisons rather than be at least as confident in $A$ as one is in $C$.

What about norms with multiple normative operators—the B formulations? Consider the Bo+ form of the transitivity norm.

> Bo+ If one is obligated to be at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is obligated to be at least as confident in $A$ as in $C$.

As I explain above, principles like this merely extend normative pressure that one is already under. One is obligated to be at least as confident in $A$ as one is in $C$, if one is also obligated to be at least as confident in $A$ as one is in $B$ and at least as confident in $B$ as one is in $C$. However, if one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, but one is not obligated to be, then the norm is silent. As MacFarlane (MS, p. 12) points out, we want norms that allow us to make negative appraisals of agents who have attitudes they shouldn't; we don't just want to be able to extend normative pressure agents are already under.

The wide scope formulation, `Wo+`, doesn't suffer from these flaws.

`Wo+` One is obligated to see to it that if one is least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is at least as confident in $A$ as one is in $C$.

If one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then the norm applies whether or not one ought to hold those comparisons. It tells one to revise one's comparisons or embrace the consequence that one is obligated to be at least as confident in $A$ as in $C$. The main short-coming of wide scope formulations is that they don't give us determinate guidance about which attitudes to hold. However, we shouldn't expect coherence norms to provide such guidance on their own in most cases. So, hereafter, I will focus on the wide scope formulations.

### 5.3.2 Which Operator?

Now that I've ruled out the narrow scope `C` and `B` formulations in favor of wide scope `W` formulations, I'll argue that we should favor evaluative `g` formulations rather than deontic `o` and `p` formulations and reasons `r` formulations. First I'll consider the deontic norms, beginning with permissions.

I find two main problems with permission formulations. First, as MacFarlane (2004, p. 10) points out, permission formulations do not *constrain* our degrees of belief. We want probabilist norms that tell us what our degrees of belief should be like. So, permissions alone are not suitable for the formulation of our probabilist norms.

Second, coherence considerations alone cannot grant permissions about how to invest our confidence. We must also consult our evidence and, perhaps, other mitigating factors, like our limited cognitive abilities. Consider the `Wp+` form of the tautological certainty norm.

`Wp+` If $\vDash A$, then one is permitted to be certain of $A$.

As I explain above, permissions like this are strict. This norm tells us that certainty in a tautology is always permissible regardless of whatever else may be the case. However, as I note above, there may be cases in which one's evidence, including perhaps knowledge of one's own cognitive limitations, might make it impermissible for one to be certain of some tautologies. If the tautology in question is extremely complex, it's hard to imagine a scenario in which complete certainty in it is permissible.

Whereas permission formulations don't constrain our degrees of belief enough on their own, obligation formulations are often—if not always—too demanding. Consider again the `Wo+` formulation of the tautological certainty norm.

`Wo+` If $\vDash A$, one is obligated to be certain of $A$.

This norm requires logical omniscience, which is far beyond the abilities of humans. Thus, it fails to satisfy the "ought"-implies-"can" constraint on obligations. To make it less demanding, we could add a knowledge condition as in the `Wo+k` formulation, or an attitude condition as in the `Wo+a` formulation.

`Wo+k` If one knows that $\vDash A$, one is obligated to be certain of $A$.

`Wo+a` If $\vDash A$, then one is obligated to see to it that, if one has a degree of belief in $A$, one is certain of it.

Both formulations are considerably less demanding than the original `Wo+` formulation. The `Wo+k` form imposes an obligation for certainty conditional on one's logical knowledge. However, depending on the connections between knowledge, belief, and degree of belief, the `Wo+k` form could be trivial, as I noted above. The `Wo+a` formulation requires logical infallibility rather than logical omniscience. It weakens the requirement significantly without invoking the difficult concept of knowledge. However, even this weaker demand is perhaps still too strong for humans. We couldn't reasonably expect to be certain of all of the logical truths we have attitudes towards. In another sense, however, the addition of knowledge and attitude conditions may weaken the demands of the norm too much. Such formulations permit what MacFarlane calls "logical obtuseness" (MS, p. 12). For there may be simple tautologies that one is obligated to be certain of regardless of whether one knows them as such or whether one has degrees of belief towards them.

Even if these formulations are sufficiently weak to pass the "ought"-implies-"can" test without swinging too far in the other direction to permit logical obtuseness, my second objection to permission formulations applies equally to obligation formulations. Namely, coherence considerations alone cannot impose obligations about how we should invest our confidence. Considerations of evidence and our practical and cognitive limitations must also be taken into account.[152] Once again, I have in mind cases in which an agent is uncertain in a complex tautology. Just as these cases seem to show that it is

[152] I assume here that there cannot be conflicting epistemic obligations—tragic scenarios in which, say, separate evidential and coherence obligations make incompatible demands. MacFarlane questions this assumption (MS, p. 14).

not always permissible to be certain of some tautologies, they also show that we are not always obligated to be certain of tautologies.

What about reasons formulations? Because they are slack, reasons formulations do not fall afoul of demandingness concerns like obligation formulations. Consider the `Wr+` form of the tautological certainty norm.

`Wr+` If $\vDash A$, one has reason to be certain of $A$.

One can satisfy this norm even if one is uncertain of some tautologies, provided that one has overriding reasons to be uncertain. This norm doesn't *demand* logical omniscience; it merely *recommends* it.

As Broome (1999, p. 404) and MacFarlane (MS, p. 12) point out, however, this advantage of reasons formulations is also their main flaw. For we cannot appeal to slack reasons norms in order to account for the appearance that in some cases, agents' degrees of belief violate strict constraints. For example, imagine that I know that $A$ is a tautology, but I am certain that it is false. In this scenario, something has gone very wrong with my degree of belief in $A$. However, on its own, the `Wr+` formulation of the tautological certainty norm cannot indict my certainty that $A$ is false.

Perhaps this is how it should be. I criticized the deontic formulations on the basis that obligations and permissions do not depend on coherence considerations alone, but also on our evidence and limitations. So, perhaps it is an advantage of reasons formulations that they do not overstate the normative contribution of coherence concerns by stating strict constraints. Ultimately, I think that reasons formulations provide a useful language in which to describe how different epistemic considerations—

coherence, evidence, limitations, etc.—interact in all-things-considered normative appraisals of our doxastic attitudes. However, for the purposes of probabilism, we want to be able to provide strict appraisals of the coherence of agents' degrees of belief. Since the reasons operator is slack, reasons formulations are not best suited toward this purpose.

### 5.3.3 Which Evaluative Formulation?

This leaves us with the wide scope evaluative formulations, the `Wg` formulations. Ultimately, I think we want a hierarchy of higher and lower standards that allow us to make fine-grained evaluations of the coherence of agents' degrees of belief. Within this evaluative hierarchy, the `Wg+` norms provide the stricter standard, whereas the `Wg+a` norms provide a standard closer to the abilities of real agents like us. Both standards play important roles, as we will now see.

All of the `Wg` formulations avoid the problems of the obligation, permission, and reason formulations. As I explain above, the evaluative norms are not constrained by our abilities in the same way obligations are. So, they are not subject to the same kinds of demandingness concerns as the obligation formulations. Furthermore, evaluative norms characterize what is good relative to some perspective or set of values. So, we can use evaluative norms to characterize what is good in the way of degree of belief from the perspective of coherence without reference to considerations of evidence and cognitive limitations. So, while coherence considerations do not impose obligations and permissions on their own, they do support univocal evaluations. Additionally, the goodness operator is strict (relative to some set of values). So, unlike the reasons formulations, evaluative norms allow us to make strict appraisals of agents' degrees of belief from the perspective of coherence.

Which of the wide scope evaluative formulations is best? As I explain above, the negative polarity formulations are equivalent to the attitude conditional formulations except in the case of certainty norms, like the tautological certainty norm. In these cases, the positive and negative polarity formulations are equivalent, whereas the attitude conditional norms succeed in weakening the demands of the norm. Compare the `Wg+`, `Wg−`, and `Wg+a` forms of the tautological certainty norm.

`Wg+` If $\vDash A$, it is good for one to be certain of $A$.

`Wg−` If $\vDash A$, it is good for one not to be uncertain of $A$.

`Wg+a` If $\vDash A$, it is good for one to see to it that, if one has a degree of belief in $A$, one is certain of it.

Hereafter, I will set aside the negative polarity formulations in favor of the attitude conditional formulations, which better capture the spirit of the notion of coherent extendability.

Adding knowledge conditions to our norms lowers the standards these norms set when they are conditional on particular logical facts, like the tautological certainty and entailment norms.

`Wg+` If $\vDash A$, it is good for one to be certain of $A$.

`Wg+k` If one knows that $\vDash A$, it is good for one to be certain of $A$.

`Wg+` If $A \vDash B$, it is good for one to be at least as confident in $B$ as one is in $A$.

`Wg+k` If one knows that $A \vDash B$, it is good for one to be at least as confident in $B$ as one is in $A$.

While evaluative norms are not subject to "ought"-implies-"can" constraints like obligations, it still may be useful to have probabilist norms that specify lower standards. After all, the Wg+ forms of the tautological certainty and entailment norms set very high standards.

However, the Wg+k formulations are not suitable as the sole probabilist norm formulations for two main reasons. First, adding knowledge conditions to our norms restricts their applicability too much. Probabilism is intended to characterize a general notion of coherence for degrees of belief—not a notion of coherence-given-logical-knowledge. Furthermore, the knowledge conditional formulations permit a kind of logical obtuseness, as I explain above. If $A$ is a tautology, but one doesn't know it as such, the Wg+k form of the tautological certainty norm doesn't provide any constraints on one's degree of belief in $A$. Second, the analysis of knowledge is notoriously controversial, and the connections between knowledge, categorical belief, and degree of belief are unclear. So, it would likely be counterproductive to attempt to characterize degree of belief coherence in terms of logical knowledge. For these reasons, I will also set aside the knowledge conditional formulations.

That leaves us with either the Wg+ or the Wg+a formulations. The main difference between these formulations is the demandingness of the evaluative standards they set. The Wg+ forms set a very high evaluative standard of what I'll call "full-blown" coherence. Together, they recommend that one should have an extremely richly structured, opinionated, logically omniscient degree of belief set. Thus, they characterize a coherence ideal that is beyond the reach of real human agents. The Wg+a norms, on the other hand, set a less-demanding standard of what I call "coherent extendability." They don't require one to adopt any particular additional degrees of belief. Rather, they merely impose coherence on the degrees of belief one already has. If it were psychologically possible, one could satisfy these norms by having no degrees of belief at all. Together, the `Wg+a` norms make more precise the notion of extendability to probabilistic coherence mentioned in Chapters 2 and 3.[153]

Compare, for example, the `Wg+` and `Wg+a` formulations of the tautological certainty and transitivity norms.

`Wg+` If $\vDash A$, it is good for one to be certain of $A$.

`Wg+a` If $\vDash A$, it is good for one to see to it that, if one has a degree of belief in $A$, one is certain of it.

`Wg+` It is good for one to see to it that if one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is at least as confident in $A$ as in $C$.

`Wg+a` It is good for one to see to it that if one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then, if one compares $A$ and $C$, one is at least as confident in $A$ as in $C$.

Notice, as I mention above, that the `Wg+` formulations entail the `Wg+a` formulations. The `Wg+` tautological certainty norm sets the high standard of logical omniscience. It tells one to adopt the attitude of certainty in every logical truth. The `Wg+a` form, on the other hand, sets the lower—but still quite high—standard of logical infallibility. It tells

153 Note that the extendability form of the probabilist norm mentioned in Chs. 2 & 3 is not quite the same as Hawthorne's (2009) notion of proper extendability described in Ch. 4. I'll say more about this connection below in §5.4.

one to be certain of every tautology toward which one has an attitude: never be uncertain in any tautology that you consider—no mistakes!

The `Wg+` transitivity norm recommends that if one is at least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one should be at least as confident in $A$ as in $C$ or one should revise one's comparisons of $A$ and $B$ or $B$ and $C$. It requires one to extend one's degree of belief set to full-blown coherence (or revise one's current attitudes). The `Wg+a` form, on the other hand, doesn't recommend that one should be at least as confident in $A$ as in $C$ (or revise) unless one compares $A$ and $C$. Admittedly, it would be a bit strange to compare $A$ and $B$ and $B$ and $C$ but not $A$ and $C$, but it is in principle possible. So, again, the attitude conditional form of the norm doesn't tell one to adopt new attitudes; it simply imposes coherence on the attitudes one already has.

Ultimately the choice between these formulations comes down to the question of whether we want to set the high standard of full-blown coherence, or the somewhat lower standard of coherent extendability. There are a couple of reasons to be wary of setting the higher standard. First, while the `Wg+` norms aren't constrained by "ought"-implies-"can" considerations, there's another sense in which the standard of full-blown coherence is too demanding. If we adopt the `Wg+` norms as our lone set of probabilist norms, they won't allow us to make very fine-grained, informative assessments of limited agents like ourselves. For, as I've noted above, humans cannot achieve full-blown coherence. So, for example, if you managed to achieve logical infallibility, the `Wg+` tautological certainty norm would be silent on your remarkable achievement. It would simply indicate that your degree of belief set falls short of ideal coherence. So, we might be wary of adopting the `Wg+` norms in favor of less demanding standards that apply better to limited agents like ourselves.

Second, we might be wary of adopting the `Wg+` formulations because we might doubt the value of full-blown coherence over and above the standard of coherent extendability. In particular, we might doubt that being logically omniscient is more valuable than being logically infallible. And we might doubt the value of the extremely rich form of opinionation that the `Wg+` norms require: one should compare every claim in the logical closure of one's ken, and one's comparisons should be extremely fine-grained. So, in general, we might doubt the value of having the extremely rich degree of belief set required by full-blown coherence. Why should one have such a rich doxastic state unless one is prompted to do so by one's practical circumstances? Why should one be certain of extremely complex tautologies unless they impinge on one's practical life somehow? These worries are versions of the cognitive clutter objections described in Chapter 3.

I think that if we recall the purposes of probabilism I identified in Chapter 4, we can assuage these concerns. I identified three main purposes of probabilism: to characterize a notion of coherence for degrees of belief, to formulate coherence norms for degrees of belief, and to apply these norms to generate verdicts about the degrees of belief of real agents in particular cases. I think the aims of characterizing a notion of coherence for degrees of belief and formulating norms that codify this notion have priority. We can then apply these norms to generate verdicts about particular cases. We can also then attempt to reconcile coherence considerations with other relevant considerations like evidential norms and facts about our practical and cognitive limitations. I contend that the first step in in this project is to characterize a coherence ideal.

I grant that we want to characterize a notion of coherence that applies to limited agents like us. Thus, I think that we should not throw out the `Wg+a` norms. After all, since the `Wg+` norms entail them, we get them for free. And I do think that we should attempt to formulate additional less demanding evaluative standards that are within closer reach of real agents, which the `Wg+a` norms provide. But even the lower standard of coherent extendability is likely beyond the reach of real humans. Ultimately, I think we want a hierarchy of higher and lower standards that allow us to make fine-grained evaluations of the coherence of agents' degrees of belief. I think the first step in characterizing this hierarchy, however, is the specification of the highest coherence ideal, the `Wg+` norms (without attitude conditions). Even if specifying this ideal doesn't pave the way for the formulation of lower coherence standards, the ideal still has an important normative role to play, as Hawthorne (2009, p. 57-58) points out. For even if we succeed in characterizing lower coherence standards without reference to the ideal, we will still want to know how far from the ideal these standards are. So, henceforth I'll focus on this highest standard, the simple `Wg+` norms.

I also grant that probabilists want to characterize other kinds of degree of belief norms—deontic norms, hypological norms,[154] and prescriptions that allow us to make

[154] Hypological norms characterize which doxastic states are praiseworthy and blameworthy. The term "hypological" is due to Zimmerman (2002). See also Zimmerman (2006).

other kinds of normative appraisals of agents' degrees of belief in addition to evaluations of goodness and badness. And I think that probabilists want to characterize degree of belief norms from a wider perspective, encompassing additional considerations like evidential norms and facts about our practical and cognitive limitations. That is, I grant that probabilists want to characterize a wider notion of degree of belief rationality. I see this wider project as the project of subjective Bayesianism in general. However, I think that before we can tackle this more ambitious project, we need to engage in the more narrowly circumscribed project of spelling out probabilism's details, characterizing the notion of ideal coherence for degrees of belief, and then specifying lower standards more applicable to real agents. Then we can try to reconcile this ideal with other considerations, and incorporate probabilism into a wider theory of degree of belief rationality and epistemology generally.

Thus, I think we should take the cognitive clutter concerns mentioned above seriously. But I think these concerns arise from viewing probabilism as a complete (synchronic) theory of degree of belief rationality in and of itself. I, on the other hand, think that probabilism is better viewed as an account of the evaluative coherence ideal for degrees of belief. When we attempt to incorporate probabilism into a wider theory of degree of belief rationality, we will need to reconcile this ideal with considerations about evidence, and cognitive and practical limitations, including clutter concerns.

### 5.4 The `Wg+` Norms & the Ideal of Full-blown Coherence

Now that I've argued for the importance of the strong (non-attitude-conditional) `Wg+` norm formulations, I want to present in more detail the notion of full-blown coherence that they codify. So, in this section, I will present the `Wg+` translations of CCF's axioms into norms. I will also present a couple of normative translations of more intuitive theorems. Finally, I will comment on the overall coherence standard that these norms set, and place it in contrast to other coherence standards and the wider theory of degree of belief rationality.

### 5.4.1 The Wg+ Norms

Without further ado, here are the `Wg+` translations of CCF's axioms into norms.

| | |
|---|---|
| Certainty-Confidence-Connection | It is good for one to see to it that one is certain that $A$ if and only if one is at least as confident that $A$ as that $(A \vee \neg A)$. |
| Non-triviality | It's good for one to be more confident in $(A \vee \neg A)$ than $\neg(A \vee \neg A)$. |
| Minimality | It's good for one to be at least as confident in any claim $B$ as one is in $\neg(A \vee \neg A)$. |
| Completeness | It's good for one to be at least as confident in $A$ as in $B$ or at least as confident in $B$ as in $A$. |
| Transitivity | It's good for one to see to it that if one is least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is at least as confident in $A$ as in $C$. |
| Right Equivalence | It's good for one to see to it that if one is certain that $C$ and $D$ are materially equivalent and one is at least as confident in $A$ as in $C$, then one is at least as confident in $A$ as in $D$. |
| Left Equivalence | It's good for one to see to it that if one is certain that $C$ and $D$ are materially equivalent and one is at least as confident in $C$ as in $B$, then one is at least as in $D$ as in $B$. |
| Subtractivity | It's good for one to see to it that if one is certain that some claim $E$ is incompatible with $A$ and $B$ respectively and one is at least as confident in $A$-or-$E$ as in $B$-or-$E$, then one is at least as confident in $A$ as in $B$. |
| Additivity | It's good for one to see to it that if one is at least as confident in $A$ as in $B$, then if one is certain that $G$ is incompatible with $A$ and $B$ respectively, then one is at least as confident that $A$-or-$G$ as that $B$-or-$G$. |

| | |
|---|---|
| Tautological Certainty | If $\vDash A$, it's good for one to be certain of $A$. |
| Separating Equi-plausible Partitions | It's good for one to see to it that if one is more confident in $A$ than $B$, then there is an equiplausible partition $\{S_1, \dots, S_n\}$ such that for each claim, $S_k$, in this partition, one is more confident in $A$ than in the disjunction $B$-or-$S_k$. |

At the risk of being tedious, I will walk through the norms and some of their intuitive consequences, to explain the kind of coherence constraint each norm imposes. Then I will call attention to some of the features of the general standard of full-blown coherence that they jointly impose, and compare it to other coherence standards we might also be interested in characterizing.

#### 5.4.1.0 Certainty-Confidence-Connection

As its name indicates, the certainty-confidence-connection norm establishes a normative connection between certainty and confidence comparisons.

| | |
|---|---|
| Certainty-Confidence-Connection | It is good for one to see to it that one is certain that $A$ if and only if one is at least as confident that $A$ as that $(A \vee \neg A)$. |

In particular, it says that it is good for one to see to it that, if one is certain of a claim $A$, then one is at least as confident that $A$ as that the tautology $(A \vee \neg A)$ is true, and if one is at least as confident that $A$ as that $(A \vee \neg A)$, then one is certain that $A$. While the tautological certainty norm is most clearly a logical omniscience requirement among CCF's norms, the certainty-confidence-connection norm is also a kind of logical omniscience requirement. For it ties certainty to the notion of tautology. On its own, the certainty-confidence-connection norm ties certainty to the simple tautologous form $(A \vee \neg A)$. The tautological certainty norm and the other norms extend this connection to apply to all tautologies.

#### 5.4.1.1 Non-triviality

When you compare CCF's axioms with the norms listed in the preceding section, you'll notice that the formulation of the non-triviality norm is not a direct `Wg+` translation of the non-triviality axiom.

| | |
|---|---|
| Non-triviality Axiom | $\neg(A \vee \neg A) \not\succcurlyeq (A \vee \neg A)$ |
| Non-triviality Norm (`Wg+`) | It's good for one to see to it that one is more confident in $(A \vee \neg A)$ than $\neg(A \vee \neg A)$. |

This is because the non-triviality axiom is stated with negative polarity: it says that sentences of the simple contradictory form, $\neg(A \vee \neg A)$, are *never* ranked as high or higher than sentences of the simple tautologous form $(A \vee \neg A)$. This axiom is stated with negative polarity because CCF's axioms are all stated in terms of the weak comparative confidence relation, $\succcurlyeq$. However, in conjunction with the minimality axiom and the definition of $\succ$ in terms of $\succcurlyeq$, the non-triviality axiom entails what I'll call the "strict non-triviality theorem."

Strict Non-triviality Theorem $\quad (A \vee \neg A) \succ \neg(A \vee \neg A)$

The non-triviality norm above is a `Wg+` translation of this theorem. In conjunction with the other norms, this non-triviality norm tells one to be strictly more confident of each tautology than each contradiction.

Like the certainty-confidence-connection norm, the non-triviality norm is also a kind of logical omniscience requirement. It tells one to be strictly more confident in all tautologies of the simple form, $(A \vee \neg A)$, than all contradictions of the simple form, $\neg(A \vee \neg A)$. Thus, all on its own, it requires one to have attitudes towards a huge set of tautologies. As we will see, the other norms add to the demandingness of probabilism's logical omniscience requirement, but, on its own, the non-triviality axiom already sets a standard that may be beyond the reach of real agents. After all, even though the tautologous form, $(A \vee \neg A)$, and the contradictory form, $\neg(A \vee \neg A)$, have a simple basic logical forms, claims of these forms could be extremely complex—too complex for real agents to decipher given our practical and cognitive limitations. So, while the `Wg+` non-triviality norm sets a plausible evaluative coherence ideal, we would want a less demanding norm to apply to real agents. However, the non-triviality axiom can be weakened without loss to CCF. One can replace it with an axiom that simply says this, for some specific simple atomic sentence $E$: $\neg(E \vee \neg E) \not\succcurlyeq (E \vee \neg E)$. From this axiom, together with the other axioms, one can prove the more general version. Thus, the relevant norm can be stated as follows, using some specific atomic sentence $E$:

It's good for one to be more confident in $(E \vee \neg E)$ than $\neg(E \vee \neg E)$.

Such a norm would be more appropriate for the purposes of developing a more real-agent-friendly evaluative standard.

**5.4.1.2 Minimality**

The minimality norm says exactly what it seems to.

Minimality It's good for one to be at least as confident in any claim $B$ as one is in $\neg(A \vee \neg A)$.

It's good for one to be at least as confident of each claim $B$ as one is in every claim of the simple contradictory form $\neg(A \vee \neg A)$. It's worth pointing out that some of the claims that take the simple contradictory form $\neg(A \vee \neg A)$ are extremely complex. So, this norm is more demanding than it may seem at first glance. Furthermore, in conjunction with the other norms, it entails the more general and even more demanding norm that one should be at least as confident in every claim as one is in each contradiction. Thus, the minimality norm also contributes to probabilism's logical omniscience requirement. Similarly, it also contributes to the probabilist opinionation requirement, for it requires comparisons between every claim and claims that take the simple contradictory form, $\neg(A \vee \neg A)$. And, like the non-triviality norm, while the minimality norm sets a plausible ideal coherence standard, it may be beyond the reach of real agents.

**5.4.1.3 Completeness**

The completeness norm says that it's good for one to have some confidence comparison for each pair of claims.

Completeness It's good for one to be at least as confident in $A$ as in $B$ or at least as confident in $B$ as in $A$.

Thus, it recommends a very strong form of opinionation. It doesn't merely recommend that one have *some* attitude toward each claim—that is, that each claim shows up somewhere in one's confidence ordering. Rather, it recommends that one be able to compare each claim with every other. For example, it's good for you (after reading this sentence, perhaps) to have a determinate confidence comparison between the claim that the OU Women's Gymnastics team will win the 2016 NCAA National Championship and the claim that Goldbach's Conjecture is true. Like each of the other norms considered so far, the completeness norm sets a standard that real agents can't live up to due to both practical and cognitive limitations, and due to the sometimes incomplete, imprecise, or equivocal nature of our evidence.

While CCF's high normative standard of full-blown coherence demands completeness, the weaker standard of coherent extendability demands mere reflexivity and completability. So, again, for the purpose of characterizing a more real-agent-friendly coherence standard, full-blown coherence may be too demanding, and the standard of coherent extendability may be more appropriate.

#### 5.4.1.4 Transitivity

Since I have used the transitivity norm as an example throughout §§5.2-5.3, I hope that its meaning is pretty clear by now.

Transitivity — It's good for one to see to it that if one is least as confident in $A$ as in $B$ and at least as confident in $B$ as in $C$, then one is at least as confident in $A$ as in $C$.

However, I'll repeat that the `Wg+` formulation of the norm tells us that if one is at least as confident in $A$ as in $B$, and at least as confident in $B$ as in $C$, it's good for one to be at least as confident in $A$ as in $C$ or to revise one's comparisons of $A$ and $B$, and $B$ and $C$. This norm seems eminently plausible as a coherence ideal. There may be cases[155] in which our evidence or our practical or cognitive limitations seem to permit us to violate the transitivity norm, but from a perspective of coherence alone, it seems hard to deny.

#### 5.4.1.5 Equivalence

The equivalence norms say that it's good for one to see to it that if one is certain that two claims are materially equivalent, then all of one's confidence comparisons that include them agree.

[155] For some putative examples, see Fishburn (1986, p. 6).

| | |
|---|---|
| Right Equivalence | It's good for one to see to it that if one is certain that $C$ and $D$ are materially equivalent and one is at least as confident in $A$ as in $C$, then one is at least as confident in $A$ as in $D$. |
| Left Equivalence | It's good for one to see to it that if one is certain that $C$ and $D$ are materially equivalent and one is at least as confident in $C$ as in $B$, then one is at least as in $D$ as in $B$. |

So, for example, the equivalence norms say together that it's good for you to see to it that if you are certain that $C$ and $D$ are materially equivalent, then you are equally confident in them. The equivalence norms make good sense as coherence ideals, though they may be hard for real agents to satisfy in combination with the strong opinionation requirements imposed by the other norms.

#### 5.1.4.6 Additivity & Subtractivity

Like the equivalence norms, the subtractivity and additivity norms are somewhat complicated, but they mean just what they say.

| | |
|---|---|
| Subtractivity | It's good for one to see to it that if one is certain that some claim $E$ is incompatible with $A$ and with $B$, respectively, and one is at least as confident in $A$-or-$E$ as one is in $B$-or-$E$, then one is at least as confident in $A$ as in $B$. |
| Additivity | It's good for one to see to it that if one is at least as confident in $A$ as in $B$, then if one is certain that $G$ is incompatible with $A$ and $B$ respectively, then one is at least as confident that $A$-or-$G$ as that $B$-or-$G$. |

Together, they say that the addition or subtraction of (subjectively) incompatible disjuncts should not make a difference in confidence comparisons from a perspective of coherence. So, if you're certain that $E$ is incompatible with $A$ and $B$ respectively, then it's good for you to see to it that you revise these certainties or you are at least as confident in $A$ as in $B$ if and only if you are at least as confident in $A$-or-$E$ as in $B$-or-$E$. While these norms seem plausible as characterizations of the coherence ideal, there is evidence that real agents systematically violate them in the face of incomplete evidence in the Ellsberg paradox.[156]

#### 5.4.1.7 Tautological Certainty

The `Wg+` form of the tautological certainty norm should be quite familiar now.

Tautological Certainty    If $\vDash A$, it's good for one to be certain of $A$.

It makes one of the most significant contributions to probabilism's logical omniscience requirement. As we've seen, real agents can't satisfy the logical omniscience requirement, but nonetheless, the tautological certainty norm and the other contributors to the requirement play an important role in providing the coherence scaffolding of the full-blown coherence ideal.

#### 5.4.1.8 Separating Equi-plausible Partitions

The separating equi-plausible partitions norm imposes a fineness of grain on one's confidence comparisons.

| | |
|---|---|
| Separating Equi-plausible Partitions | It's good for one to see to it that if one is more confident in $A$ than $B$, then there is an equi-plausible partition $\{S_1, \ldots, S_n\}$ such that for each claim, $S_k$, in this partition, one is more confident in $A$ than in the disjunction $B$-or-$S_k$. |

It says that it's good for one to see to it that whenever one is strictly more confident in one claim $A$ than another claim $B$, there is some equiplausible partition $\{S_1, \ldots, S_n\}$—a (large) set of mutually incompatible, jointly exhaustive claims in all of which one is equally confident—such that one is more confident in $A$ than $B$ or $S_k$. Thus, if you're

[156] See Ellsberg (1961) for the Ellsberg Paradox and see Fishburn (1986) for an interpretation of the results as a counterexample to the additivity and subtractivity norms.

strictly more confident in $A$ than $B$, the norm tells you to revise this comparison, or it tells you that your confidence structure should have two features. First, your confidence structure should include equi-plausible partitions. Second, the claims within these partitions should separate $A$ and $B$ in the sense that you should be more confident in $A$ than $B$-or-$S_k$ and more confident in $B$-or-$S_k$ than $B$. If you satisfy this norm, in effect, whenever you're strictly more confident in $A$ than $B$, there's a degree—a degree as precise as the equi-plausible partition is large, and as small as your confidence in each claim in the partition—to which you are more confident in $A$ than $B$.

Real agents may not satisfy this norm both because of their practical and cognitive limitations, and because their evidence may preclude the precision it requires. They may not have equi-plausible partitions that give their degree of belief set enough structure to make fine-grained comparisons. Or, even if they do have equi-plausible partitions, they may find it difficult to satisfy the condition for all claims in combination with the completeness norm. While real agents won't satisfy the norm, it nonetheless characterizes a coherence ideal. An ideally coherent degree of belief set has a fine-grained structure.

For the purposes of characterizing the more real-agent-friendly coherence standard of coherent extendability, we could replace the separating equi-plausible partitions norm with what we might call the "proper extendability norm."

| | |
|---|---|
| Proper Extendability | It's good for one to see to it that, in principle, one's confidence comparisons can be extended so as to: (i) definitely compare each pair of claims (i.e., leave no pair of claims confidence-incomparable); and (ii) (to see to it that) if one is more confident in $A$ than $B$, then (it's possible, in principle, that) one can incorporate into one's comparisons an equi-plausible partition (e.g. a fair lottery of size n) |

consisting of claims $\{S_1, \ldots, S_n\}$ such that for each claim, $S_k$, in this partition, one is more confident in $A$ than in the disjunction $B$ or $S_k$.

Rather than recommending that agents actually have complete confidence comparisons and separating equi-plausible partitions, the proper extendability norm merely recommends that agents' degrees of belief should be coherently augmentable with additional attitudes.

#### 5.4.1.9 The Entailment Norm & Other Intuitive Norms

In addition to the norms generated by translating the axioms, there are infinitely many additional norms generated by translating the theorems that follow from the axioms. I've mentioned some of them informally above, but I want to call attention once more to the entailment norm in particular.

| | |
|---|---|
| Entailment Norm | If $A \vDash B$, then it's good for one to be at least as confident in $B$ as in $A$. |

The entailment norm is very straightforward and intuitive. It extends the logical omniscience requirement imposed by the non-triviality, minimality, and tautological certainty norms and it is often very intuitive to apply to real cases.

I should note that it is well-documented that humans regularly violate this norm. Kahneman & Tversky's (1983) conjunction fallacy is one such common example. However, the fact that humans systematically violate the entailment norm doesn't undermine its status as a coherence ideal.

### 5.4.2 Full-blown Coherence

Together, the `Wg+` norms set a very high standard of coherence. According to this ideal, a fully coherent degree of belief set is very richly structured, opinionated, and logically omniscient. As we've seen, the logical omniscience requirement extends far beyond mere certainty in all tautologies. It also consists of being strictly more confident in all tautologies than all contradictions, equally confident in logically equivalent claims, at least as confident in all claims as in the claims that entail them, and so on. Attitudes toward logical truths and claims that bear entailment relations thus provide the scaffolding of ideally coherent degree of belief sets. The extent of the opinionation of an ideally coherent degree of belief set is not just that the ideally coherent agent has some degree of belief in each claim. Additionally, this opinionation also consists in having a determinate confidence comparison for every pair of claims, including claims that concern vastly disparate matters. Indeed, not only does the ideally coherent agent have such comparisons, but they are extremely fine-grained. So, not only can the ideally coherent agent compare all claims, but they can compare them with an arbitrarily fine degree of precision.

Above, I have pointed out various ways that real agents can't live up to this coherence ideal. Due to our practical and cognitive limitations, we can't have so many attitudes, we can't be expected to recognize all tautologies and entailment relations as such, etc. I have also pointed out various ways that it may be inappropriate for us to attempt to satisfy this ideal given our evidential, practical, and cognitive limitations. For all of these reasons, the `Wg+` norms and the standard of full-blown coherence that they codify may not be especially useful in the appraisal and instruction of real human agents. However, as I explained in §5.3, this should not be seen as a mark against the `Wg+` norms or the value of full-blown coherence ideal. Instead, it should be viewed as reason to additionally articulate lower coherence standards, which are less demanding and which may be more easily reconciled with considerations about our evidential, practical, and cognitive limitations. The `Wg+a` norms, along with the proper extendability norm, are good places to start articulating such lower, real-agent-friendly standards.

**5.5 Justifying the Norms**

In the previous section, my primary aim was to present the `Wg+` norms and to characterize the ideal of full-blown coherence that they encode. However, I also tried to show how each norm provides a reasonable coherence constraint on degrees of belief, and I explained how apparent counterexamples to these norms weigh coherence considerations against considerations of evidence and limitations. Thus, these apparent counterexamples do not undermine the norms as statements of what is good in the way of confidence from a perspective of coherence alone. In this sense, I've provided some justification for the `Wg+` norms and CCF, generally.

However, as we saw in Chapter 2, there is a long tradition among probabilists of providing "Arguments for Probabilism" of a few well-known kinds—the Dutch book arguments, representation theorem arguments, and epistemic utility arguments. I haven't yet committed to one of these arguments to justify probabilism. So, you might be wondering, which kind of argument I will endorse. Will I opt for a version of one of the standard arguments, or will I provide some new kind of justification? The answer is that I'm not going to provide a traditional justification of probabilism, and I'm also not going to venture to provide some completely new kind of argument either. In this section, I'll explain why I don't endorse any of the standard justifications for probabilist norms, but why I also think it is nonetheless reasonable to be committed to probabilism given the intuitiveness of its norms and consequences, and its successes in illuminating problems in epistemology, confirmation theory, and decision theory.

Before I explain why I don't endorse any of the traditional justifications for probabilism, I want to make a few comments about the phenomenon of Arguments for Probabilism. For there's something odd about these arguments. Epistemologists don't usually justify their analyses and norms with arguments quite like the traditional arguments for probabilism. Usually, I think, epistemologists try to do what I've done above, motivate the norm, analysis, or principle in question and defend it against counterexamples.

So, why do probabilists feel compelled to offer these grand arguments for their norms? I think there are a few reasons. First, probabilism grew out of the degree of belief analysis of *probability*. So, early probabilists like Ramsey and de Finetti gave a lot of attention to the question of why degrees of belief should conform to the probability axioms. The traditional arguments for probabilism are intended to demonstrate the close connection between *probability* and *degree of belief*. Second, as I've noted many times, probabilism's details are usually left unstated, so the normative connections between the probability axioms and degrees of belief can be a little unclear. So, as I explained in Chapter 2, the arguments for probabilism serve as much to fill in probabilism's details as they do to justify it. They serve to characterize degrees of belief, and they help flesh out what it means for degrees of belief to conform to the probability axioms.

Before I explain the particular reasons why I reject each of the major arguments as justifications for probabilism, I want to explain generally why I won't undertake the project of giving a capital "A" Argument for Probabilism in this dissertation. First, as I explained in Chapter 2 and again in Chapter 4, I think that we can get a firm handle on the notion *degree of belief* without providing an operational definition, analysis, or explication of it. So, I've assumed primitivism about *degree of belief* as a working assumption. I welcome a successful analysis or explication, but, following Eriksson & Hájek (2007), I don't think that any of the current options is completely successful. So, I don't think the traditional arguments for probabilism are called for in order to analyze and legitimatize *degree of belief*. Likewise, I think we can characterize a notion of degree of belief coherence without providing a Dutch book argument, representation theorem argument, or epistemic utility argument. Second, as I just explained, I think that the traditional arguments for probabilism serve as much (or more) to characterize probabilism than they serve to justify it, and I think we can undertake this project of characterization without engaging in the project of justification. Finally, as I will explain below, there are aspects of each of the traditional arguments that make me leery of whole-heartedly endorsing any one of them as a successful justification for probabilism. Instead, I think that each of the traditional arguments contributes to the project of characterizing degrees of belief and degree of belief coherence, but none of them fully justifies the probabilist coherence norm. To use Hájek's term, I think the traditional arguments help us to "triangulate" to probabilism (2008, p. 816).

So, why in particular do I reject the traditional arguments as justification of probabilism? I won't consider each argument in detail, but I will present the main reasons I don't endorse any one of them as *the* justification for probabilism. The main reason I don't endorse a Dutch book argument for probabilism is that I reject the betting quotient analysis of degrees of belief. I think that betting quotients are often a good indication of degrees of belief, but they do not constitute degrees of belief, and they indicate or measure degrees of belief only imperfectly. So, I reject versions of the argument that draw a tight connection between degrees of belief and betting quotients. I am sympathetic to versions of the argument that draw looser connections between degrees of belief and betting quotients. I have in mind here Christensen's (1996, 2004) depragmatized versions of the argument. However, I also think that these versions provide, as Titelbaum (2013, p. 286) puts it, at best an "indirect" justification for probabilistic degree of belief coherence. Thus, I see these arguments more as illustrations or dramatizations of the importance of degree of belief coherence than I see them as conclusive arguments for probabilism.

I don't fully endorse representation theorem arguments for similar reasons.[157] Just as I deny the tight connection between degrees of belief and betting quotients, I also deny the tight connection between degrees of belief and preferences. Setting aside the

[157] I wish to distinguish, once again, the traditional Representation Theorem Arguments for probabilism, which tie degrees of belief to preferences, from appeals to representation theorems that demonstrate the representability of comparative confidence functions with quantitative probability functions.

problems with the representation theorem argument itself,[158] I reject the analysis or explication of *degree of belief* in terms of *preference*.

My reasons for avoiding commitment to an epistemic utility argument for probabilism are a little different. My primary reasons for avoiding the Dutch book and representation theorem arguments have to do with what I see as inaccurate characterizations of degrees of belief. Since the epistemic utility arguments don't seek so much to analyze *degree of belief*, I have different reasons for being leery of them. First, I avoid calibration arguments because I am not sure that calibration—matching relative frequencies—is a genuine epistemic good, or that the goodness of coherence stems from the goodness of calibration, if indeed it is good. My reasons for avoiding the gradational accuracy arguments are different still. I acknowledge that accuracy is a genuine epistemic good, and I am sympathetic to the idea that coherence is good because of its connection to accuracy across possible worlds. However, I don't provide a gradational accuracy argument for probabilism in this dissertation because I think that the accuracy measures that these arguments rely on are every bit as controversial as probabilism itself.[159] So, while I think that accuracy arguments provide a useful tool to help explore the connections between coherence and accuracy, and to suss out the normative assumptions that underlie probabilism, I don't think that these arguments provide genuine justification for probabilism at present.

[158] See Zynda (2000), Hájek (2008), and Meacham & Weisberg (2011) for discussion.

[159] Thanks to Jim Hawthorne on this point.

So, how do I think probabilism is justified? First, I want to note again that I would welcome a decisive Argument for Probabilism, if there were one. But since I think each of the major candidates has significant problems, I think we can find at least some alternative justification for probabilism in three main ways. First, I think we can justify probabilism by giving the kind of intuitive justification and defense of its norms that I attempt to provide in the previous section. Second, I think, following Hájek (2008), that we may be able to appeal to revised—weakened—versions of the traditional arguments for probabilism to triangulate to a characterization and justification of probabilism. Finally, following Eriksson & Hájek (2007) and Hájek (2008), I think that we can lend some justification to probabilism via a "proof in the pudding" strategy. That is, we can justify probabilism by appealing to its many successes in illuminating problems in epistemology, confirmation theory, and decision theory. Since probabilism provides such intuitive, helpful analyses of problem cases in these areas, it stands to reason that there is at least a kernel of truth in it.

**5.6 Conclusion**

In this chapter, I set out to continue the project of stating a unified, complete probabilist modeling framework by generating a set of norms from CCF's confidence axioms. Above, I explained that we need to carefully formulate these norms in order to make probabilism's content and consequences clear. I presented the options for formulating norms from CCF's axioms, and I argued for the `Wg+` formulations as a characterization of the evaluative ideal of coherence for degrees of belief. I also mentioned some possibilities for lowering this standard of coherence to reconcile coherence considerations with considerations about evidence, and our practical and cognitive limitations. Finally, I explained why I don't endorse any of the typical arguments for probabilism to justify the notion of full-blown coherence I have characterized.

# 6 Conclusion

I've had three main aims in this dissertation. First, I wanted to show why probabilists need to better articulate the foundations of their view. Second, I wanted to characterize what exactly it means to provide well-articulated foundations for probabilism. Finally, I wanted to provide an account of the core aspects of probabilism's foundations. I set out to accomplish the first task in Chapters 2 and 3, and I set out to accomplish the latter two tasks in Chapters 4 and 5.

In Chapter 2, I presented a typical statement of probabilism's core tenets as well as the most prominent (and widely discussed) options for filling in the details. I presented probabilists' characterizations of degrees of belief, and their justifications for positing degrees of belief. I also presented the probability calculus, its interpretation as a source of degree of belief norms, and the standard arguments probabilists use to justify the probabilist norm that degrees of belief should satisfy the axioms of probability theory. I presented this material primarily in order to acquaint the reader with probabilism, but also in order to begin to demonstrate the need for probabilists to fill in the details of their view.

In Chapter 3, I developed in earnest the argument that probabilists must articulate unified foundations for their view. I presented the central issues in the dialectic between probabilists and their critics, and I drew out some of the major themes from this exchange. I presented three main sources of objections to probabilism, and I presented several of the major responses to these objections. I showed how the objections I surveyed fall into three main kinds: descriptive adequacy objections, demandingness objections, and value objections. I also showed how the probabilist responses to these objections tend to be concessive or dogmatic in the face of these concerns. I showed how this dialectic raises various questions about probabilism's methodological and normative foundations, and how probabilist responses to objections begin to reveal different answers to these foundational questions. I concluded with an argument that probabilists must articulate unified foundations for their view. I argued that while probabilists have a variety of responses to the main objections to their view (both concessive and dogmatic), several of these responses are in tension. They are not clearly compatible. So, I concluded that probabilists must articulate the foundations of their view in a unified, systematic way in order to clarify their view's content and aims, and in order to provide a fully developed rebuttal of the common objections to their view.

In the remainder of the dissertation, I set out to begin the project of spelling out unified foundations for probabilism. I started with probabilism's methodological foundations in Chapter 4, where I provided an account of probabilism's aims, its formalism, and its modeling methodology. I provided an account of the nature and aims of formal modeling in general and probabilism in particular. I concluded that probabilism has three main aims: characterizing certain philosophical concepts (*degree of belief*, *coherence*, and *rationality*), formulating norms for degrees of belief, and generating normative verdicts about degrees of belief in particular cases. I also adopted and presented Titelbaum's (2013) account of the elements of a well-articulated modeling framework. Finally, I set out to characterize the elements of a well-articulated probabilist modeling framework, CCF (the Comparative Confidence Framework), beginning with what I call its methodological foundations. In particular, I set out CCF in terms of a logical (rather than set-theoretic) formal system, and in terms of a fundamentally qualitative, comparative (rather than quantitative) representation of degrees of belief.

In Chapter 5, I provided an account of probabilism's normative foundations. In particular, I defended an account of the translation of CCF's axioms and theorems into degree of belief norms. I explained that we need to carefully formulate these norms in order to make probabilism's content and consequences clear. I presented many options for formulating norms from CCF's axioms, and I argued for what I called the `Wg+` formulations, which characterize degree of belief coherence as an evaluative ideal, which I called "full-blown coherence." Finally, I explained why I don't endorse any of the typical arguments to justify probabilism.

The result of my efforts is CCF. CCF is a logical, fundamentally qualitative and comparative probabilist modeling framework. It is intended primarily to characterize a notion of ideal degree of belief coherence, which I call "full-blown coherence." And it provides norms associated with this coherence standard, which can be applied to evaluate the coherence of the degrees of belief of real agents in particular cases.

CCF avoids concerns about the excessively precise numerical representation of degrees of belief that plague traditional probabilism. Indeed, it clarifies when the numerical representation of degrees of belief is truly appropriate. Furthermore, it is well suited to the representation of the degrees of belief of non-ideal agents, and agents whose evidence is incomplete, imprecise, or equivocal.

CCF is also sensitive to concerns about the demandingness of its normative standard of full-blown coherence. As I explained in Chapter 5, full-blown coherence is an extremely high normative standard—well beyond the reach of real human agents, with all of our cognitive and practical limitations. Though CCF characterizes a coherence ideal that is beyond the reach of real agents, it is just one part of the wider subjective Bayesian theory of doxastic and pragmatic rationality. Many of the apparent tensions that emerge in probabilist responses to objections are due to the fact that probabilism's purposes are not clear. Is probabilism supposed to characterize mere coherence norms for degrees of belief? Or is it supposed to codify all of the synchronic degree of belief norms, encompassing both coherence considerations and evidential considerations? Is probabilism an epistemic ideal or is it supposed to provide norms that guide real agents, with all of our practical, computational, and mnemonic limitations? CCF characterizes probabilism as a coherence ideal. A complete statement of this wider theory will include an account of how to reconcile coherence considerations with evidential norms and considerations about human limitations. CCF is a first step in a more careful articulation of the wider subjectivist Bayesian account of doxastic and practical rationality.

## References


Adams, R. (1974). Theories of actuality. *Noûs*, *8*, 211-231.

Anscombe, G.E.M. (1957). *Intention*. Oxford: Basil Blackwell.

Armstrong, D. (1989). *A combinatorial theory of possibility*. Cambridge.

Arntzenius, F. (2003). Some problems for conditionalization and reflection. *Journal of Philosophy, 100*(7), 356-370.

Bjerring, J.C. (2010). *Non-ideal epistemic spaces*. Doctoral Dissertation, Australian National University.

Bostrom, N. (2007). Sleeping Beauty and self-location: A hybrid model. *Synthese, 157*(7), 59-78.

Bradley, D. (2012). Four problems about self-locating belief. *Philosophical Review, 121*(2), 149-177.

Broome, J. (1999). Normative requirements. *Ratio, 12*(4), 398-419.

Capotorti, A. & Vantaggi, B. (2000). Axiomatic characterization of partial ordinal relations. *International Journal of Approximate Reasoning, 24*, 207-219.

Chalmers, D. (2011). Frege's puzzle and the objects of credence. *Mind, 120*, 587-635.

Chalmers, D. (2011a). The nature of epistemic space. In A. Egan & B. Weatherson (eds.) *Epistemic modality* (pp. 60-107). New York: Oxford.

Cherniak, C. (1986). *Minimal rationality*. Cambridge: MIT.

Christensen, D. (1996). Dutch-book arguments depragmatized: Epistemic consistency for partial believers*. The Journal of Philosophy, 93*(9), 450-479.

Christensen, D. (2001). Preference-based arguments for probabilism. *Philosophy of Science, 68*(3), 356-376.

Christensen, D. (2004). *Putting logic in its place: Formal constraints on rational belief*. New York: Oxford.

De Finetti, B. (1931). Sul significato soggettivo della probabilità. *Fundamenta Matematicae, 17*, 293-329.

De Finetti, B. (1964). Foresight: Its logical laws, its subjective sources. In H. Kyburg & H. Smokler (eds.) *Studies in subjective probability* (pp. 93-158). New York: Wiley.

De Finetti, B. (1974). *Theory of probability.* New York: Wiley.

Earman, J. (1992). *Bayes or bust? A critical examination of Bayesian confirmation theory*. Cambridge: MIT.

Easwaran, K. (2011). Bayesianism I: Introduction and arguments in favor. *Philosophy Compass, 6*, 321-332.

Easwaran, K. (2011a). Bayesianism II: Applications and criticisms. *Philosophy Compass, 6*, 312-320.

Easwaran, K., & Fitelson, B. (2015). Accuracy, coherence, and evidence. *Oxford Studies in Epistemology*, *5*, 61-96.

Eells, E. (1985). Problems of old evidence. *Pacific Philosophical Quarterly, 66*, 283-302.

Elga, A. (2000). Self-locating belief and the Sleeping Beauty problem. *Analysis, 60*(266), 143-147.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75*, 643-669.

Eriksson, L., & Hájek, A. (2007). What are degrees of belief? *Studia Logica, 86*, 183-212.

Field, H. (2009). What is the normative role of logic? *Proceedings of the Aristotelian Society*, supplementary volume LXXXIII, 251-268.

Fine, K. (1977). Postscript. In A. Prior (Ed.), *Worlds, times, and selves* (pp. 116-161). Amherst: UMASS.

Fine, T. (1973). *Theories of probability*. New York: Academic.

Fishburn, P. (1986). The axioms of subjective probability. *Statistical Science, 1*(3), 335-358.

Fitelson, B. (2008). A decision procedure for probability calculus with applications. *Review of Symbolic Logic, 1*, 111-125.

Fodor, J. (1975). *The language of thought*. New York: Cromwell.

Foley, R. (2009). Belief, degrees of belief, and the Lockean thesis. In F. Huber & C. Schmidt-Perti (Ed.), *Degrees of belief* (pp. 47-74). New York: Springer.

Frankish, K. (2009). Partial belief and flat-out belief. In F. Huber & C. Schmidt-Perti (Ed.), *Degrees of belief* (pp. 75-93). New York: Springer.

Frege, G. (1997). On *Sinn* and *Bedeutung*. In M. Beany (ed.) *The Frege reader* (pp. 151-180). New York: Blackwell.

Gilovich, T., Griffin, D., & Kahneman, D. (2002*). Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge.

Goldman, A. (1986). *Epistemology and cognition*. Cambridge: Harvard.

Hacking, I. (1967). Slightly more realistic personal probability. *Philosophy of Science, 34*(4), 311-325.

Hájek, A. (2003). What conditional probability could not be. *Synthese, 137*, 273-323.

Hájek, A. (2008). Arguments for—or against—probabilism? *British Journal for the Philosophy of Science, 59*, 793-819.

Hájek, A. (2009). Dutch book arguments. In P. Anand, P. Pattanaik, & C. Puppe (eds.) *The handbook of rational and social choice* (pp. 173-192). New York: Oxford.

Hájek, A. (2013). Is strict coherence coherent? *Dialectica, 66*(3), 411-424.

Hájek, A. (MSa). A puzzle about partial belief.

Hájek, A. (MSb). Staying regular?

Hajek, A. (MSc). Formal methods and Bayesian epistemology: Two Hegelian dialectics.

Hájek, A., & Hartmann, S. (2010). Bayesian epistemology. In J. Dancy (Ed.), *A companion to epistemology* (pp. 93-105). Oxford: Blackwell.

Hansson, S. O. (2000). Formalization in philosophy. *The Bulletin of Symbolic Logic, 6*(2), 162-175

Harman, G. (1984). Logic and reasoning. *Synthese, 60*, 107-127.

Harman, G. (1986). *Change in view*. Cambridge: MIT.

Hawthorne, J. (2009). The Lockean thesis and the logic of belief. In F. Huber & C. Schmidt-Perti (Ed.), *Degrees of belief* (pp. 49-74). New York: Springer.

Hawthorne, J. (2011). Bayesian confirmation theory. In S. French & J. Saatsi (eds.) *The continuum companion to the philosophy of science* (pp. 197-213). Continuum Press.

Hawthorne, J. (MS). A logic of comparative support: qualitative conditional probability relations representable by Popper functions. Forthcoming in A. Hajek & c. Hitchcock (eds.) *The Oxford handbook of probability and philosophy*. Oxford UP.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach* (3rd ed.). La Salle: Open Court.

Huber, F. (2009). Belief and degrees of belief. In F. Huber & C. Schmidt-Perti (Ed.), *Degrees of belief* (pp. 1-33). New York: Springer.

Huber, F. (2014). Formal representations of belief. Retrieved from http://plato.stanford.edu/archives/spr2013/entries/formal-belief/

Humberstone, I.L. (1992). Direction of fit. *Mind, 101*(401), 59-83.

Jaynes, E. (2003). *Probability theory: The logic of science*. New York: Cambridge.

Jeffrey, R. (1965). *The logic of decision*. Chicago. U. of Chicago Press.

Jeffrey, R. (1970). Dracula meets Wolfman: Acceptance vs. partial belief. In M. Swain (ed.) *Induction, acceptance, and rational belief* (pp. 157-185). New York: Springer.

Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science, 65*(4), 575-603.

Joyce, J. (1999). *The foundations of causal decision theory*. New York: Cambridge.

Joyce, J. (2004). Bayesianism. In P. Rawling & A.R. Mele (eds.), *The Oxford handbook of rationality* (pp. 132-155). New York: Oxford.

Joyce, J. (2005). How probabilities reflect evidence. *Philosophical Perspectives, 19*, 154-178.

Joyce, J. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber & C. Schmidt-Perti (Eds.), *Degrees of belief* (pp. 263-297). New York: Springer.

Joyce, J. (2010). A defense of imprecise credences. *Philosophical Perspectives, 24*, 281-323.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, and Giroux.

Kahneman, D, Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New Yourk: Cambridge.

Kahneman, D, & Tversky, A. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293-315.

Kaplan, M. (2002). Decision theory and epistemology. In P.K. Moser (ed.) *The Oxford handbook of epistemology* (pp. 434-457). New York: Oxford.

Kaplan, M. (2010). Modest probabilism. *Synthese, 176*, 41-55.

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

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

King, J. C. (2012). Structured propositions. Retrieved from http://plato.stanford.edu/archives/win2012/entries/propositions-structured/

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

Kyburg, H. (1961). *Probability and the logic of rational belief*. Middletown: Wesleyan UP.

Lehman, R. (1955). On confirmation and rational betting. *Journal of Symbolic Logic, 20*, 251-262.

Levinstein, B. (2013). *Accuracy as epistemic utility*. Doctoral Dissertation, Rutgers University, New Brunswick.

Lewis, D. (1974). Radical interpretation. *Synthese, 27*, 331-344.

Lewis, D. (1979). Attitudes *de dicto* and *de se*. *The Philosophical Review, 88*(4), 513-543.

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. Jeffrey (Ed.), *Studies in inductive logic and probability* (vol. 2). Los Angeles: University of California Press.

Lewis, D. (1986). *On the plurality of worlds*. New York: Oxford.

MacFarlane, J. (2000). *What does it mean to say that logic is formal?* Doctoral Dissertation, University of Pittsburgh.

MacFarlane, J. (MS). In what sense (if any) is logic normative for thought? Retrieved from: http://johnmacfarlane.net/normativity_of_logic.pdf

Maher, P. (1993). *Betting on theories*. New York: Cambridge.

Makinson, P. (1965). The paradox of the preface. *Analysis, 25*(6), 205-207.

McGrath, M. (2012). Propositions. Retrieved from http://plato.stanford.edu/archives/sum2012/entries/propositions/

Meacham, C. (2008). Sleeping Beauty and the dynamics of *de se* belief. *Philosophical Studies, 138*(2), 245-269.

Meacham, C. & Weisberg, J. (2011). Representation theorems and the foundations of decision theory. *Australasian Journal of Philosophy, 89*(4), 641-663.

Menzel, C. (2008). Possible worlds. Retrieved from http://plato.stanford.edu/archives/win2013/entries/possible-worlds/

Pettigrew, R. (2011). Epistemic utility arguments for probabilism. Retrieved from http://plato.stanford.edu/archives/win2013/entries/epistemic-utility/

Plantinga, A. (1974). *The nature of necessity*. New York: Oxford.

Plantinga, A. (1976). Actualism and possible worlds. *Theoria, 42*, 139-160.

Pohl, R. (2012). *Cognitive illusions: A handbook on fallacies and biases in thinking*. Malden: Blackwell.

Popper, K. (1955). Two autonomous axiom systems for the calculus of probabilities. *British Journal for the Philosophy of Science*, 6(21), 51-57.

Quine, W. (1969). Propositional objects. In *Ontological relativity and other essays*. New York: Columbia.

Ramsey, F. P. (1964). Truth and probability. In H. Kyburg & H. Smokler (Eds.) *Studies in subjective probability* (pp. 61-92). New York: Wiley.

Rényi, A. (1955). On a new axiomatic theory of probability. *Acta mathematca Hungarica*, 6(3-4), 285-335.

Rényi, A. (1970). *Foundations of probability*. San Francisco: Holden-Day.

Roberts, F. (1985). *Measurement theory*. New York: Cambridge.

Savage, L.J. (1967). Difficulties in the theory of personal probability. *Philosophy of Science, 34*(4), 305-310.

Savage, L.J. (1972). *The foundations of statistics*. New York: Dover.

Schwitzgebel, E. (2011). Belief. Retrieved from http://plato.stanford.edu/archives/win2011/entries/belief/

Seidenfeld, T. (1985). Calibration, coherence, and scoring rules. *Philosophy of science, 52*, 274-294.

Shimony, A. (1988). An Adamite derivation of the principles of the calculus of probability. In J. Fetzer (ed.) *Probability and causality* (pp. 78-89). New York: Reidel.

Smith, M. (1987). The Humean theory of motivation. *Mind, 96,* 36-61.

Smith, M. (1988). Reason and desire. *Proceedings of the Aristotelian Society, 88*, 243-258.

Smith, M. (1994). *The moral problem*. New York: Blackwell.

Sobel, D. & Copp, D. (2001). Against direction of fit accounts of belief and desire. *Analysis, 61*(1), 44-53.

Sturgeon, S. (2008). Reason and the grain of belief. *Noûs, 42*(1), 139-165.

Sturgeon, S. (2010). Confidence and coarse grained attitudes. *Oxford Studies in Epistemology, 3*, 126-149.

Talbott, W. (2005). Review of *Putting logic in its place: Formal constraints on rational belief*. *Notre Dame Philosophical Reviews*. URL: http://ndpr.ne.edu.

Talbott, W. (2011). Bayesian epistemology. Retrieved from http://plato.stanford.edu/archives/sum2011/entries/epistemology-bayesian/

Titelbaum, M. (2013). *Quitting certainties: A Bayesian framework for modeling degrees of belief*. New York: Oxford.

Titelbaum, M. (2015). Rationality's fixed point (Or: In defense of right reason). *Oxford Studies in Epistemology, 5*, 253-294.

Titelbaum, M. (MS). Fundamentals of Bayesian Epistemology. Oxford University Press.

van Fraassen, B. (1984). Belief and the will. *Journal of Philosophy, 81*(5), 235-256.

van Fraassen, B. (1999). Conditionalization, a new argument for. *Topoi, 18*(2), 93-96.

Vineberg, S. (2011). Dutch book arguments. Retrieved from: http://plato.stanford.edu/archives/sum2011/entries/dutch-book/.

Walley, P. (1991). *Statsitical reasoning with imprecise probabilities*. London: Chapman and Hall.

Weisberg, M. (2007). Three kinds of idealization. *The Journal of Philosophy, 104*(12), 639-659.

Weisberg, J. (2011). Varieties of Bayesianism. In D. Gabbay, S. Hartmann, & J. Woods (Eds.), *Inductive Logic* (vol. 10, pp. 477-552). Waltham, MA: Elsevier.

Wheeler, G. (MS). Formal Epistemology.

White, R. (2009). Evidential symmetry and mushy credence. *Oxford Studies in Epistemology, 3*, 161-188.

Williamson, T. (2000.) *Knowledge and its limits*. New York: Oxford.

Wong, D, et al. (1991). Axiomatization of qualitative belief structure. *IEEE Transactions on Systems, Man, and Cybernetics, 21*(4), 726-732.

Zimmerman, M. (2002). Taking luck seriously. *The Journal of Philosophy, 99*(11), 553-576.

Zimmerman, M. (2006). Moral luck: A partial map. *Canadian Journal of Philosophy, 36*(4), 585-608.

Zynda, L. (1996). Coherence as ideal of rationality. *Synthese, 109*(2), 175-216.

Zynda, L. (2000). Representation theorems and realism about degrees of belief. *Philosophy of Science*, 67(1), 45-69.

## Appendix: Alternative Comparative Confidence Axioms

In this appendix, I'll present two alternative sets of axioms for CCF's rudimentary confidence relations. Both alternative sets of axioms are equivalent to CCF's rudimentary confidence axioms. I present these alternative axioms here not because they have different consequences from CCF's axioms (they don't), but for three main reasons. First, the alternative axioms are similar to other extant approaches in the literature on qualitative probability and comparative confidence. Second, the alternative axioms have some features that make them more suitable for some purposes than CCF's axioms. Third, the alternative axioms are somewhat simpler than CCF's axioms.

CCF's rudimentary confidence axioms are stated so as to avoid appealing to the notions of entailment and generic tautology, except in the tautological certainty axiom. This makes CCF's other axioms a little more real agent friendly; for, some tautologies are extremely complex and some entailment relations are extremely difficult to compute. Thus, axioms that demand knowledge of any tautology whatsoever, or any entailment relation whatsoever, may be too demanding for real agents to follow in many cases. Ultimately, both CCF's axioms and the alternative sets of axioms below have the same consequences, since they are equivalent. So, in the grand scheme, they all make the same strong demands. However, CCF's axioms themselves, taken individually, are a little easier for real agents to follow when translated into norms.[160]

[160] Excepting the tautological certainty axiom.

For the purposes of crafting a general theory of doxastic and practical rationality, real-agent-friendliness is a desirable feature of a set of comparative confidence axioms. However, for some more specific purposes, like the purpose of characterizing the highest ideal of degree of belief coherence without respect to the abilities of real agents, we might not be so squeamish about real-agent-unfriendliness. The main differences between CCF's axioms and the alternative sets of axioms below are that the alternative axioms appeal to the notions of entailment and tautology more freely, and neither set of axioms appeals to a certainty predicate. For these reasons, the alternative axioms are a little simpler than CCF's axioms. Thus, for the purposes of characterizing the highest ideal of degree of belief coherence, the alternative sets of axioms below may be preferable.

The first set of alternative axioms is due to James Hawthorne.[161] As before, a relation, $\succcurlyeq$, defined on a standard language for predicate logic $\mathcal{L}$ is a rudimentary confidence relation if and only if it satisfies the following axioms. For all $A, B, C, D \in \mathcal{L}$:

| | | |
|---|---|---|
| A1 | there are some sentences $F$ and $G$ such that $F \not\succcurlyeq G$ | (non-triviality); |
| A2 | if $\vDash A$, then $A \succcurlyeq B$ (for every sentence $B$) | (maximality); |
| A3 | $A \succcurlyeq A$ | (reflexivity); |
| A4 | if $A \succcurlyeq B$ and $B \succcurlyeq C$, then $A \succcurlyeq C$ | (transitivity); |
| A5 | if for some $E$, $\vDash \neg(A \wedge E)$, $\vDash \neg(B \wedge E)$, and $(A \vee E) \succcurlyeq (B \vee E)$, then $A \succcurlyeq B$ | (subtractivity); |

[161] In private communication.

A6 if $A \succcurlyeq B$, then for all $G$ such that $\vDash \neg(A \wedge G)$ and $\vDash \neg(B \wedge G)$, $(A \vee G) \succcurlyeq (B \vee G)$ (additivity).

In Chapter 4, I noted that CCF's comparative confidence are based on the comparative confidence axioms of Hawthorne (2009), which are in turn based on the axioms of Savage (1972). The second set of alternative axioms is a logical analogue of Capotorti & Vantaggi's (2000, pp. 210) set-theoretic comparative confidence axioms, which are similar to some early attempts at axiomatizing qualitative probability.[162] Once again, a relation, $\succcurlyeq$, defined on a standard language for predicate logic $\mathcal{L}$ is a rudimentary confidence relation if and only if it satisfies the following axioms (where $\bot$ is an arbitrary contradiction and $\top$ is an arbitrary tautology). For all $A, B, C, D \in \mathcal{L}$:

B1 $\bot \not\succcurlyeq \top$ (non-triviality);

B2 if $A \vDash B$, then $B \succcurlyeq A$ (entailment);

B3 $A \succcurlyeq A$ (reflexivity);

B4 if $A \succcurlyeq B$ and $B \succcurlyeq C$, then $A \succcurlyeq C$ (transitivity);

B5 if for some $E$, $\vDash \neg(A \wedge E)$, $\vDash \neg(B \wedge E)$, and $(A \vee E) \succcurlyeq (B \vee E)$, then $A \succcurlyeq B$ (subtractivity);

B6 if $A \succcurlyeq B$, then for all $G$ such that $\vDash \neg(A \wedge G)$ and $\vDash \neg(B \wedge G)$, $(A \vee G) \succcurlyeq (B \vee G)$ (additivity).

It's obvious that the A axioms follow from the B axioms. The converse may not be so clear. In particular, one might want a proof that axiom B2 follows from the A axioms. (It's obvious that Axiom A2 follows from Axiom B2.)

162 For example de Finetti (1931) as cited in Capotorti & Vantaggi (2000) and Fishburn (1986).

*Theorem.* From the A axioms we have the following: If $A \vDash B$, then $B \succcurlyeq A$.

*Proof.* Suppose $A \vDash B$. Then, $\vDash \big((A \wedge B) \vee \neg A\big)$, so from A2, $\big((A \wedge B) \vee \neg A\big) \succcurlyeq (A \vee \neg A)$. Notice also that $\vDash \neg\big((A \wedge B) \wedge \neg A\big)$ and $\vDash \neg(A \wedge \neg A)$, so $(A \wedge B) \succcurlyeq A$ (from A5). Similarly, $\vDash (B \vee \neg B)$, so from A2, $(B \vee \neg B) \succcurlyeq \big((A \wedge B) \vee \neg A\big)$. Notice also that $\vDash \neg(B \wedge \neg B)$ and $\vDash \neg\big((A \wedge B) \wedge \neg B\big)$, so $B \succcurlyeq (A \wedge B)$ (from A5). Then $B \succcurlyeq A$ (A4). ■

It's also worth noting that, like CCF's rudimentary confidence axioms, neither of these axiomatizations is sufficient for unique representability with numerical probability functions on its own. However, both sets are sufficient for unique probabilistic representation when they are supplemented with CCF's Axiom X, as is the case with CCF's rudimentary confidence relations.