Coherence

Branden Fitelson

with

Kenny Easwaran
&
David McCarthy

For my mother.

Contents

List of Tables

List of Figures

Preface

Sitting down to write the preface of a book about epistemic coherence requirements can be somewhat tricky (and ironic). According to the position articulated in this book, it can be rational (in cases such as these) to acknowledge the inevitability of some errors having creeped into this investigation, while maintaining belief in each of the individual claims asserted herein.

This book has three main aims (not necessarily in this order): (1) to layout a general and unified framework for grounding (formal, synchronic) epistemic coherence requirements, (2) to apply this framework to three types of judgment (belief, comparative confidence, and numerical credence), and (3) to discuss various issues and problems that arise along the way. The book is broken-up into three parts — one for each of the three types of judgment mentioned in aim (2). I recommend reading the book in the order in which it is presented. The decision to progress from belief to comparative confidence to numerical credence was intentional. And, the insights garnered at each stage should be helpful for understanding subsequent parts of the book. Each of the three parts has a "Positive Phase" chapter and a "Negative Phase" chapter. In the Positive Phase, the proposed explication of coherence will be presented. In the Negative Phase, various possible objections, shortcomings, *etc*., to the proposed explication will be discussed.

The first part of the book (on belief) draws on joint work with Kenny Easwaran, and the second part of the book (on comparative confidence) draws on joint work with David McCarthy. The third part of the book (on numerical credence) also benefitted greatly from my discussions (and collaborations) with Kenny and David. I owe both Kenny and David a huge debt of gratitude for their stalwart collaborative efforts (and for allowing me to bring our joint work together into this monograph).

I have also had the great pleasure (and honor) of presenting this material in many places around the world. Indeed, I've had many valuable discussions about this project over the past several years. So many, in fact, that it would be infeasible to list all of my wonderful interlocutors. I must, however, thank the following people (in alphabetical order) for providing particularly important feedback on the project at various stages of its development: tbd. . .

I would like to be able to say that the errors which remain in the book are all my fault. But, since I think some such errors would be inevitable regardless of the author, I will not make that assertion. Having said that, I am confident that the book is *immeasurably* better because of the generosity and skill of my many friends and colleagues. I am truly blessed.

Finally, I would like to thank my wife Tina Eliassi-Rad for her love and support over the past twenty years. It is impossible to articulate her importance. Neither this philosopher nor this book would exist were it not for her.

Part I

Belief[1]

[1]This part of the book is based on joint work with Kenny Easwaran. Note: this part needs to be re-written, since (as it stands) many parts of it are lifted directly from (Easwaran and Fitelson, 2014). Similar things can also be said for some of the later sections of this manuscript as well. That's why this is a (very) *preliminary draft*, which should not be quoted or cited (*ergo* subtle watermark).

CHAPTER 1

Belief: Positive Phase

> We put before ourselves the standard of consistency and construct these elaborate rules to ensure its observance. But . . . it is [not] even clear that consistency is always advantageous; it may well be better to be sometimes right than never right.
>
> Frank Plumpton Ramsey

1. Background

This monograph is about formal, synchronic, epistemic, coherence requirements (of ideal rationality). We begin by explaining how we will be using each of these (five) key terms.

Formal epistemic coherence requirements involve properties of judgment sets that are logical (and, in principle, determinable *a priori*). These are to be distinguished from other less formal and more substantive notions of coherence that one encounters in the epistemological literature. For instance, so-called "coherentists" like BonJour (1985) use the term in a less formal sense which implies (*e.g.*) that coherence is truth-conducive. While there will be conceptual connections between the accuracy of a doxastic state and its coherence (in the sense we have in mind), these connections will be quite weak (certainly too weak to merit the conventional honorific "truth-conducive"). All of the varieties of coherence to be discussed in Part I will be intimately related to deductive consistency. Consequently, whether a set of judgments is coherent will be determined by (*i.e.*, will supervene on) logical properties of the set of propositions that are the objects of the judgments in question (analogous remarks will apply *mutatis mutandis* to the requirements discussed in Parts II and III).

Synchronic epistemic coherence requirements apply to the doxastic states of agents at individual times. These are to be distinguished from diachronic coherence requirements (*e.g.*, conditionalization, reflection, *etc.*), which apply to sequences of doxastic states across times. In the present study, we will be concerned only with the former.[2]

Epistemic requirements are to be distinguished from, *e.g.*, pragmatic requirements. Starting with Ramsey (1928), the most well-known arguments for probabilism as a formal, synchronic, coherence requirement for credences have depended

[2] See Titelbaum (2013) for an excellent recent survey of the contemporary literature on (Bayesian) diachronic epistemic coherence requirements. Some, *e.g.*, Moss (2013) and Hedden (2013), have argued that there are no diachronic epistemic rational requirements (*i.e.*, that there are only synchronic epistemic rational requirements). We take no stand on this issue here. But, we will assume that there are (some) synchronic epistemic rational requirements of the sort we aim to explicate (see *fn.* 7).

on the pragmatic connection of belief to action. For instance, "Dutch Book" arguments and "Representation Theorem" arguments (Hájek 2008) aim to show that an agent with non-probabilistic credences (at a given time t) must (thereby) exhibit some sort of "pragmatic defect" (at t).[3] Following Joyce (1998; 2009), we will be focusing on non-pragmatic (*viz.*, epistemic) defects implied by the synchronic incoherence (in a precise, general sense to be explicated below) of an agent's doxastic state. Specifically, we will be concerned with two aspects of doxastic states that we take to be distinctively epistemic: (a) how accurate a doxastic state is, and (b) how much evidential support a doxastic state has. We will call these (a) alethic and (b) evidential aspects of doxastic states, respectively.[4]

Coherence requirements are global and wide-scope. Coherence is a global property of a judgment set in the sense that it depends on properties of entire set in a way that is not (in general) reducible to properties of individual members of the set. Coherence requirements are wide-scope in Broome's (2007) sense. They will be expressible using "shoulds" (or "oughts") that take wide-scope over some logical combination(s) of judgments. As a result, coherence requirements will not (in general[5]) require specific attitudes toward specific individual propositions. Instead, coherence requirements will require the avoidance of certain combinations of judgments. We use the term "coherence" — rather than "consistency" — because (a) the latter is typically associated with classical deductive consistency (which, as we'll see shortly, we do not accept as a necessary requirement of epistemic rationality), and (b) the former is used by probabilists when they discuss analogous requirements for degrees of belief (*viz.*, probabilism as a coherence requirement for credences). Because our general approach can be applied to many types of judgment (*e.g.*, belief, comparative confidence, and numerical credence) we prefer to maintain a common parlance for the salient requirements in all of these settings.

Finally, and most importantly, when we use the term "requirements", we are talking about necessary requirements of ideal epistemic rationality.[6] The hallmark of a necessary requirement of epistemic rationality N is that if a doxastic state S violates N, then S is (thereby) epistemically irrational. However, just because a doxastic state S satisfies a necessary requirement N, this does not imply that S is (thereby) rational. For instance, just because a doxastic state S is coherent

[3]We realize that "depragmatized" versions of these arguments have been presented (Christensen, 1996). But, even these versions of the arguments trade essentially on the pragmatic role of doxastic attitudes (in "sanctioning" bets, *etc.*). In contrast, we will only be appealing to epistemic connections of belief to truth and evidence. That is, our arguments will not explicitly rely upon any connections between belief and action.

[4]The alethic/evidential distinction is central to the pre-Ramseyan debate between James (1896) and Clifford (1877). Roughly speaking, "alethic" considerations are "Jamesian", and "evidential" considerations are "Cliffordian". We will (for the most part) be assuming for the purposes of this monograph that alethic and evidential aspects exhaust the distinctively epistemic properties of doxastic states. But, our framework could be generalized to accommodate additional dimensions of epistemic evaluation (should there be such). We'll return to this issue in the Negative Phase below, where we consider an objection along these lines.

[5]There are two notable exceptions. It will follow from our approach that (a) rational agents should never believe individual propositions ($\perp$) that are logically self-contradictory, and (b) that rational agents should never disbelieve individual propositions ($\top$) that are logically true. Similar "limiting cases" of our coherence requirements will also arise in Parts II and III.

[6]We adopt Titelbaum's (2013, chapter 2) locution "necessary requirement of (ideal) rationality" as well as (roughly) his usage of that locution (as applied to synchronic requirements).

(*i.e.*, just because S satisfies some formal, epistemic coherence requirement), this does not mean that S is (thereby) rational (as S may violate some other necessary requirement of epistemic rationality). Thus, coherence requirements in the present sense are (formal, synchronic) necessary conditions for the epistemic rationality of a doxastic state.[7] Our talk of the epistemic (ir)rationality of doxastic states is meant to be evaluative (rather than normative[8]) in nature. To be more precise, we will (for the most part) be concerned with the evaluation of doxastic states, relative to an idealized standard of epistemic rationality. Sometimes we will speak (loosely) of what agents "should" do — but this will (typically) be an evaluative sense of "should" (*e.g.*, "should on pain of occupying a doxastic state that is not ideally epistemically rational"). If a different sense of "should" is intended, we will flag this by contrasting it with the idealized/evaluative "should" that features in our rational requirements.

We will be modeling the (synchronic) doxastic states of agents in a rather simple and idealized fashion. More precisely, we will be considering agents S who make explicit judgments (at a given time) regarding each of the propositions on some finite *agenda* $\mathcal{A}$. We will not be evaluating (or considering) S's attitudes toward propositions that are not on the agenda $\mathcal{A}$. That is, we will not be discussing questions regarding the implicit commitments S might be taken to have, in light of their explicit commitments regarding the propositions in $\mathcal{A}$. Similarly, we will not be concerned with what inferences may or may not be rational or reasonable for S to make, in light of their explicit commitments on $\mathcal{A}$.

Propositions can be thought of either as finite sets of (classical) possible worlds or (alternatively) as elements of some (corresponding) finite Boolean algebra of propositions $\mathcal{B}$. Because we are working with a simple, classical possible worlds framework, we will be presupposing that our agents are logically omniscient in the following precise sense: if p and q are logically equivalent, then S makes the same judgments regarding p and q (*e.g.*, in the case of belief, this implies that either she believes both p and q or she disbelieves both of them). It is important to note that our weak logical omniscience assumption does not imply that an agent's judgments satisfy any closure conditions (*e.g.*, an agent S can be such that she believes that P and she believes that Q, but she does not believe that $P \mathbin{\&} Q$). In this sense, our framework differs from the traditional (Stalnakerian) framework for modeling belief in which this sort of simple closure condition for belief is typically assumed.[9] This is crucial for us, since we will end-up defending coherence requirements that

[7] For simplicity, we will assume that there *are* some (synchronic, epistemic) rational requirements in the first place. We are well aware of the current debates about the very existence of rational requirements (*e.g.*, coherence requirements). Specifically, we are cognizant of the debates between Kolodny (2007) and others, *e.g.*, Broome (2007). Here, we will simply adopt the non-eliminativist stance of Broome *et al.* who accept the existence of rational requirements (*e.g.*, coherence requirements). We will not try to justify our non-eliminativist stance here, as this would take us too far afield. However, as we will explain below, even coherence eliminativists like Kolodny should be able to benefit from our approach and discussion (see *fn.* 45).

[8] Normative principles support attributions of blame or praise of agents, and are (in some sense) action guiding. Evaluative principles support classifications of states (occupied by agents) as "defective" *vs* "non-defective" ("bad" *vs* "good"), relative to some evaluative standard (Smith, 2005, §3). Here, we are evaluating doxastic states relative to (alethic and evidential) epistemic standards.

[9] Our assumption of weak logical omniscience could be defended — as a requirement of ideal epistemic rationality — in various ways. See, for instance Titelbaum (2013); Stalnaker (1991, 1999). But, we will be content to adopt this assumption as a (mere) simplifying idealization.

are weaker than deductive consistency. As such, it is important for us to be able to evaluate the explicit judgments made by agents (on agendas) without assuming any constraints (*e.g.*, closure) on propositions that are not on the agenda (*e.g.*, the agent's implicit commitments).

We will use the notation $\ulcorner B(p) \urcorner$ to express the claim that S *believes* p and $\ulcorner D(p) \urcorner$ to express the claim that S *disbelieves* p. For simplicity, we will assume that, for each proposition $p \in \mathcal{A}$, S either believes p or disbelieves p (and not both). That is to say, we will be focusing on agendas $\mathcal{A}$ over which S is *opinionated* (*i.e.*, we will not be evaluating sets of judgments involving *suspension* of belief). This focus on opinionated sets of explicit commitments is (of course) not meant to imply that suspension of judgment is never rational or reasonable (Sturgeon, 2008; Friedman, 2013). Rather, opinionation is merely a(nother) simplifying idealization, which (for present purposes) will result in no significant loss of generality. We will return to the opinionation/suspension issue in our discussion of the Preface Paradox below (and also in the Negative Phase).

Here is a simple example which illustrates the kind of judgment sets that we will be evaluating. Consider a fair lottery with n tickets, exactly one of which is the winner. For each $j \leqslant n$ (for $n \geqslant 3$), let p_j be the proposition that the j^{th} ticket is not the winning ticket; let q be the proposition that some ticket is the winner; and, let these $n+1$ propositions exhaust the agenda $\mathcal{A}$. Finally, let LOTTERY be the following opinionated judgment set on $\mathcal{A}$

$$\textsc{Lottery} \overset{\text{def}}{=} \left\{B(p_j) \mid 1 \leqslant j \leqslant n\right\} \cup \{B(q)\}.$$

Any such judgment set LOTTERY will be deductively inconsistent (since q entails that one of the p_j is false). But, as we will see below, LOTTERY sets need not be evaluated as incoherent (or irrational) in our framework.

Our use of such simple and idealized (*viz.*, naïve) formal models is intentional, as our aims here are rather modest. Mainly, we aim to present the simplest version of our general framework, and to contrast our proposed coherence requirements with some alternative (equally naïve) requirements that have been defended in the philosophical literature (*e.g.*, classical deductive consistency). For present purposes, our idealizations will turn out to be fairly benign. Having said that, some potential shortcomings and problems with our naïve approach will be discussed in the Negative Phase. Now that the stage is set, it will be instructive to look at the most well-known "coherence requirement" in the intended sense of that term.

2. Deductive Consistency, Accuracy, and Evidence

The most well-known example of a formal, synchronic, epistemic coherence requirement for belief is the (putative) requirement of deductive consistency.

(CB) All agents S should (at any given time t) have sets of beliefs (*i.e.*, sets of belief-contents) that are (classically) deductively consistent.

Many philosophers have assumed that (CB) is a necessary requirement of ideal epistemic rationality. That is, many philosophers have assumed that (CB) is true, if its "should" is interpreted as "should on pain of occupying a doxastic state that is not ideally epistemically rational". Interestingly, in our perusal of the literature, we haven't been able to find many (general) arguments in favor of the claim that

(CB) is a rational requirement. One potential argument along these lines takes as its point of departure the (so-called) *Truth Norm* for belief.[10]

(TB) All agents S should (at any given time t) have beliefs that are true.[11]

Presumably, there is some sense of "should" for which (TB) comes out true, *e.g.*, "should on pain of occupying a doxastic state that is not perfectly accurate" (see *fn.* 11). But, we think most philosophers would not accept (TB) as a rational requirement.[12] Nonetheless, (TB) clearly implies (CB) — in the sense that all agents who satisfy (TB) must also satisfy (CB). So, if one violates (CB), then one must also violate (TB). Moreover, violations of (CB) are the sorts of things that one can (ideally, in principle) be in a position to detect *a priori*. Thus, one might try to argue that (CB) is a necessary requirement of ideal epistemic rationality, as follows. If one is (ideally, in principle) in a position to know *a priori* that one violates (TB), then one's doxastic state is not (ideally) epistemically rational. Therefore, (CB) is a rational requirement. While this (TB)-based argument for (CB) may have some *prima facie* plausibility, we'll argue that (CB) itself seems to be in tension with another plausible epistemic norm, which we call the *Evidential Norm* for belief.

(EB) All agents S should (at any given time t) have beliefs that are supported by the total evidence.

For now, we're being intentionally vague about what "supported" and "the total evidence" mean in (EB), but we'll precisify these locutions in due course.[13]

---

[10]We will use the term "norm" (as opposed to "requirement") to refer to local/narrow-scope epistemic constraints on belief. The Truth Norm (as well as the Evidential Norm, to be discussed below) is local in the sense that it constrains each individual belief — it requires that each proposition believed by an agent be true. This differs from the rational requirements we'll be focusing on here (*viz.*, coherence requirements), which are global/wide-scope constraints on sets of beliefs. Moreover, the sense of "should" in norms (generally) differs from the evaluative sense of "should" that we are associating with rational requirements (see *fn.* 13).

[11]Our statement of (TB) is (intentionally) somewhat vague here. Various precisifications of (TB) have been discussed in the contemporary literature. See Thomson (2008), Wedgwood (2002), Shah (2003), Gibbard (2005) and Boghossian (2003) for some recent examples. The subtle distinctions between these various renditions of (TB) will not be crucial for our purposes. For us, (TB) plays the role of determining the correctness/accuracy conditions for belief (*i.e.*, it determines the alethic ideal for belief states). In other words, the "should" in our (TB) is intended to mean something like "should on pain of occupying a doxastic state that is not entirely/perfectly correct/accurate". In this sense, the version of (TB) we have in mind here is perhaps most similar to Thomson's (2008, Ch. 7).

[12]Some philosophers maintain that justification/warrant is factive (Littlejohn 2012; Merricks 1995). In light of the Gettier problem, factivity seems plausible as a constraint on the type of justification required for knowledge (Zagzebski 1994; Dretske 2013). However, factivity is implausible as a constraint on (the type of justification required for) rational belief. As such, we assume that "is supported by the total evidence" (*i.e.*, "is justified/warranted") is not factive. This assumption is kosher here, since it cross-cuts the present debate regarding (CB). For instance, Pollock's defense of (CB) does not trade on the factivity of evidential support/warrant (see *fn.* 15). And, neither does Ryan's argument for (CB) — which we discuss in detail in the Negative Phase (Ryan, 1996).

[13]The evidential norm (EB) is [like (TB)] a local/narrow-scope principle. It constrains each individual belief, so as to require that it be supported by the evidence. We will not take a stand on the precise content of (EB) here, since we will (ultimately) only need to make use of certain (weak, probabilistic) consequences of (EB) that are implicated in the formal coherence of doxastic states. However, the "should" of (EB) is not to be confused with the "should" of (TB). It may be useful (heuristically) to read the "should" of (EB) as "should on pain of falling short of the Cliffordian ideal" and the "should" of (TB) as "should on pain of falling short of the Jamesian ideal" (see *fns.* 4 & 10).

Versions of (EB) have been endorsed by various "evidentialists" (Clifford 1877; Conee & Feldman 2004). Interestingly, the variants of (EB) we have in mind conflict with (CB) in some ("paradoxical") contexts. For instance, consider the following example, which is a "global" version of the Preface Paradox.

> **Preface Paradox**. Let **B** be the set containing all of *S*'s justified first-order beliefs. Assuming *S* is a suitably interesting inquirer, this set **B** will be a very rich and complex set of judgments. And, because *S* is fallible, it is reasonable to believe that some of *S*'s first-order evidence (or putative first-order evidence) will (inevitably) be misleading. As a result, it seems reasonable to believe that some beliefs in **B** are false. Indeed, we think *S* herself could be justified in believing this very second-order claim. But, of course, adding this second-order belief to **B** renders *S*'s overall doxastic (belief) state deductively inconsistent.

We take it that, in (some) such preface cases, an agent's doxastic state may satisfy (EB) while violating (CB). Moreover, we think that (some) such states need not be (ideally) epistemically irrational. That is, we think our Preface Paradox (and other similar examples to be discussed below) establish the following key claim:

(†) (EB) does not entail (CB). [*i.e.*, the Evidential Norm does not entail that deductive consistency is a requirement of ideal epistemic rationality.]

It is beyond the scope of this monograph to provide a thorough defense of (†).[14] Foley (1992) sketches the following, general "master argument" in support of (†).

> ... if the avoidance of recognizable inconsistency were an absolute prerequisite of rational belief, we could not rationally believe each member of a set of propositions and also rationally believe of this set that at least one of its members is false. But this in turn pressures us to be unduly cautious. It pressures us to believe only those propositions that are certain or at least close to certain for us, since otherwise we are likely to have reasons to believe that at least one of these propositions is false. At first glance, the requirement that we avoid recognizable inconsistency seems little enough to ask in the name of rationality. It asks only that we avoid certain error. It turns out, however, that this is far too much to ask.

We think Foley is onto something important here. As we'll see, Foley's argument dovetails nicely with our approach to grounding coherence requirements for belief.

So far, we've been assuming that agents facing Prefaces (and similar paradoxes of deductive consistency) may be opinionated regarding the (inconsistent) sets of propositions in question (*i.e.*, that the agents in question either believe or disbelieve each proposition on the salient agendas). In the next section, we consider the

[14] Presently, we are content to take (†) as a *datum*. However, definitively establishing (†) requires only the presentation of one example (preface or otherwise) in which (CB) is violated, (EB) is satisfied, and the doxastic state in question is not (ideally) epistemically irrational. We think our Preface Paradoxes suffice. Be that as it may, we think Christensen (2004), Foley (1992), and Klein (1985) have given compelling reasons to accept (†). And, we'll briefly parry some recent philosophical resistance to (†) below. One might even want to strengthen (†) so as to imply that satisfying (EB) sometimes requires the violation of (CB). Indeed, this stronger claim is arguably established by our Preface Paradox cases. In any event, we will, in the interest of simplicity, stick with our weaker rendition of (†).

possibility that the appropriate response to the Preface Paradox (and other similar paradoxes) is to suspend judgment on (some or all) propositions implicated in the inconsistency.

3. Suspension of Judgment to the Rescue?

Some authors maintain that opinionation is to blame for the discomfort of the Preface Paradox (and should be abandoned in response to it). We are not moved by this line of response to the Preface Paradox. We will now briefly critique two types of "suspension strategies" that we have encountered.

It would seem that John Pollock (1983) was the pioneer of the "suspension strategy". According to Pollock, whenever one recognizes that one's beliefs are inconsistent, this leads to the "collective defeat" of (some or all of) the judgments comprising the inconsistent set. That is, the evidential support that one has for (some or all of) the beliefs in an inconsistent set is defeated as a result of the recognition of said inconsistency. If Pollock were right about this (in full generality), it would follow that if the total evidence supports each of one's beliefs, then one's belief set must be deductively consistent. In other words, Pollock's general theory of evidential support (or "warrant"[15]) must entail that (†) is false. Unfortunately, however, Pollock does not offer much in the way of a general argument against (†). His general remarks tend to be along the following lines (Pollock 1990, *p*. 231).[16]

> The set of warranted propositions must be deductively consistent. ... If a contradiction could be derived from it, then reasoning from some warranted propositions would lead to the denial (and hence defeat) of other warranted propositions, in which case they would not be warranted.

The basic idea here seems to be that, if one (knowingly) has an inconsistent set of (justified) beliefs, then one can "deduce a contradiction" from this set, and then "use this contradiction" to perform a "*reductio*" of (some or all of) one's (justified) beliefs.[17] Needless to say, anyone who is already convinced that (†) is true will find this general argument against (†) unconvincing. Presumably, anyone who finds themselves in the midst of a situation that they take to be a counterexample to (EB)

[15]Pollock uses the term "warranted" rather than "supported by the total evidence". But, for the purposes of our discussion of Pollock's views, we will assume that these are equivalent. This is kosher, since, for Pollock, "S is warranted in believing p" means "S could become justified in believing p through (ideal) reasoning proceeding exclusively from the propositions he is objectively justified in believing" (Pollock 1990, *p*. 87). Our agents, like Pollock's, are "idealized reasoners", so we may stipulate (for the purposes of our discussion of Pollock's views) that when we say "supported by the total evidence", we just mean whatever Pollock means by "warranted". Some (Merricks 1995) have argued that Pollock's notion of warrant is factive (see *fn.* 12). This seems wrong to us (in the present context). If warrant (in the relevant sense) were factive, then Pollock wouldn't need such complicated responses to the paradoxes of consistency — they would be trivially ruled out, *a fortiori*. This is why, for present purposes, we interpret Pollock as claiming only that (EB) entails the consistency of (warranted) belief sets [(CB)], but not necessarily the truth of each (warranted) belief [(TB)].

[16]The ellipsis in our quotation contains the following parenthetical remark: "It is assumed here that an epistemic basis must be consistent." That is, Pollock gives no argument(s) for the claim that "epistemic bases" (*viz.*, Pollockian "input propositions") must be consistent.

[17]Ryan (1991; 1996) gives a similar argument against (†). See the Negative Phase for a detailed discussion of Ryan's argument for against (†) [and for (CB)]. Nelkin (2000) endorses Ryan's argument, as applied to defusing the lottery paradox as a counterexample to (EB) ⇒ (CB).

⇒ (CB) should be reluctant to perform "*reductios*" of the sort Pollock seems to have in mind, since it appears that consistency is not required by their evidence. Here, Pollock seems to be assuming a closure condition (*e.g.*, that "is supported by the total evidence" is closed under logical consequence/competent deduction) to provide a *reductio* of (†). It seems clear to us that those who accept (†) would/should reject closure conditions of this sort. We view (some) Preface cases as counterexamples to both consistency and closure of rational belief.[18]

While Pollock doesn't offer much of a general argument for ¬(†), he does address two apparent counterexamples to ¬(†): the lottery paradox and the preface paradox. Pollock (1983) first applied this "collective defeat" strategy to the lottery paradox. He later recognized (Pollock 1986) that the "collective defeat" strategy is far more difficult to (plausibly) apply in the case of the Preface Paradox. Indeed, we find it implausible on its face that the propositions of the (global) Preface "jointly defeat one another" in any probative sense. More generally, we find Pollock's treatment of the Preface Paradox quite puzzling and unpersuasive.[19] Be that as it may, it's difficult to see how this sort of "collective defeat" argument could serve to justify ¬(†) in full generality. What would it take for a theory of evidential support to entail ¬(†) — in full generality — *via* a Pollock-style "collective defeat" argument? We're not sure. But, we are confident that any explication of "supported by the total evidence" (or "warranted"/"justified") which embraces a phenomenon of "collective defeat" that is robust enough to entail the falsity of (†) will also have some undesirable (even unacceptable) epistemological consequences.[20]

We often hear another line of response to the Preface that is similar to (but somewhat less ambitious than) Pollock's "collective defeat" approach. This line of response claims that there is something "heterogenous" about the evidence in the Preface Paradox, and that this "evidential heterogeneity" somehow undermines the claim that one should believe all of the propositions that comprise the Preface Paradox. The idea seems to be[21] that the evidence one has for the first-order beliefs (in **B**) is a (radically) different kind of evidence than the evidence one has for the second-order belief (*i.e.*, the belief that renders **B** inconsistent in the end). And, because these bodies of first-order and second-order evidence are so heterogenous, there is no single body of evidence that supports both the first-order beliefs and the second-order belief in the Preface case. So, believing all the propositions of the

[18]See (Steinberger 2014) for a thorough analysis of the consequences of the preface paradox for principles of deductive reasoning [*i.e.*, "bridge principles" *à la* (MacFarlane 2004)]. And, see our discussion of Ryan's argument (in the Negative Phase) for additional insight into which specific assumptions about evidential support (or justification) we think are to blame for the failure of closure.

[19]We don't have the space here to analyze Pollock's approach to the Preface Paradox. Fortunately, Christensen (2004) has already done a very good job of explaining why "suspension strategies" like Pollock's can not, ultimately, furnish compelling responses to the Preface.

[20] For instance, it seems to us that any such approach will have to imply that "supported by the total evidence" is (generally) closed under logical consequence (or competent deduction), even under complicated entailments with many premises. See (Korb 1992) for discussion regarding (this and other) unpalatable consequences of Pollockian "collective defeat".

[21]We've actually not been able to find this exact line of response to the Preface anywhere in print, but we have heard this kind of line defended in various discussions and Q&A's. The closest line of response we've seen in print is Leitgeb's (2013) approach, which appeals to the "heterogeneity" of the subject matter of the claims involved in the Preface. This doesn't exactly fall under our "evidential heterogeneity" rubric, but it is similar enough to be undermined by our Homogeneous Preface case. See, also, (Leitgeb, 2014).

Preface is not, in fact, the epistemically rational thing to do.[22] Hence, the apparent tension between (EB) and (CB) is merely apparent.

We think this line of response is unsuccessful, for three reasons. First, can't we just gather up the first-order and second-order evidential propositions, and put them all into one big collection of total Preface evidence? And, if we do so, why wouldn't the total Preface evidence support both the first-order beliefs and the second-order belief in the Preface case? Second, we only need one Preface case in which (EB) and (CB) do genuinely come into conflict in order to establish (†). And, it seems to us that there are "homogeneous" versions of the Preface which do not exhibit this (alleged) kind of "evidential heterogeneity". Here's one such example.

> **Homogeneous Preface Paradox**. John is an excellent empirical scientist. He has devoted his entire (long and esteemed) scientific career to gathering and assessing the evidence that is relevant to the following first-order, empirical hypothesis: (H) all scientific/empirical books of sufficient complexity contain at least one false claim. By the end of his career, John is ready to publish his masterpiece, which is an exhaustive, encyclopedic, 15-volume (scientific/empirical) book which aims to summarize (all) the evidence that contemporary empirical science takes to be relevant to H. John sits down to write the Preface to his masterpiece. Rather than reflecting on his own fallibility, John simply reflects on the contents of (the main text of) his book, which constitutes very strong inductive evidence in favor of H. On this basis, John (inductively) infers H. But, John also believes each of the individual claims asserted in the main text of the book. Thus, because John believes (indeed, knows) that his masterpiece instantiates the antecedent of H, the (total) set of John's (rational) beliefs is inconsistent.

In our Homogeneous Preface, there seems to be no "evidential heterogeneity" available to undermine the evidential support of John's ultimate doxastic state. Moreover, there seems to be no "collective defeat" looming here either. John is simply being a good empirical scientist (and a good inductive non-skeptic) here, by (rationally) inferring H from the total, H-relevant inductive scientific/empirical evidence.[23] It is true that it was John himself who gathered (and analyzed, *etc.*) all of this inductive evidence and included it in one hugely complex scientific/empirical book. But, we fail to see how this fact does anything to undermine the (ideal) epistemic rationality of John's (ultimate) doxastic state. So, we conclude that the "heterogeneity strategy" is not an adequate response to the Preface.[24] More generally,

---

[22]Presumably, then, the rational thing to do is suspend judgment on some of the Preface propositions. But, which ones? As in the case of Pollock's "suspension strategy", it remains unclear to us precisely which propositions fail to be supported by the total evidence (and why).

[23]One might worry that, because evidence is factive (Williamson, 2000), John is not basing his belief in H on the total, H-relevant *evidence* (assuming H is true). However, even if evidence is factive and H is true, John's inference to H could be supported by the total H-relevant evidence. To see this, note that John may accept the following claim: *some vast majority* of the claims in my book is true, *i.e.*, some vast majority of the claims in my book constitutes the total H-relevant evidence (Leitgeb, 2014). And, if each such vast majority of claims inductively supports H, then (as far as we can see) John's inference to H may well be rational *and* based on the total H-relevant evidence — no matter what that collection of total H-relevant evidence happens to be.

[24]We said we rejected the "heterogenous evidence" line of response to the Preface for three reasons. Our third reason is similar to the final worry we expressed above regarding Pollock's "collective defeat" strategy. We don't see how a "heterogeneity strategy" could serve to establish ¬(†) in full generality, without presupposing something implausible about the nature of evidential support, *e.g.*, that support is preserved by competent deduction (see *fn.* 20).

we think our Homogeneous Preface case undermines any strategy that maintains one should never believe all the propositions in any Preface.[25]

We maintain that (adequate) responses to the Preface Paradox need not require suspension of judgment on (any of) the Preface propositions. Consequently, we would like to see a (principled) response to the Preface Paradox (and other paradoxes of consistency) that allows for (full) opinionation with respect to the propositions in the Preface agenda. Indeed, we will provide just such a response (to all paradoxes of consistency) below.

Before presenting our framework (and response), we will compare and contrast our own view regarding the Preface Paradox (and other paradoxes of consistency) with the views recently expressed by a pair of philosophers who share our commitment to (†) — *i.e.*, to the claim that Preface cases (and other similar cases) show that deductive consistency is not a necessary requirement of ideal epistemic rationality.

4. Christensen and Kolodny on Coherence Requirements

We are not alone in our view that Prefaces (and other paradoxes of deductive consistency) suffice to establish (†).[26] For instance, David Christensen and Niko Kolodny agree with us about Prefaces and (†). But, Christensen and Kolodny react to the paradoxes of deductive consistency in a more radical way. They endorse

($\star$) There are no coherence requirements (in the relevant sense) for belief.

That is to say, both Christensen and Kolodny endorse eliminativism regarding all (formal, synchronic, epistemic) coherence requirements for belief. It is illuminating to compare and contrast the views of Christensen and Kolodny with our own views about paradoxes of consistency and proper responses to them.

Christensen (2004) accepts the following package of pertinent views.[27]

(C_1) Partial beliefs (*viz.*, credences) are subject to a formal, synchronic, epistemic coherence requirement (of ideal rationality): probabilism.

($\star$) beliefs are not subject to any formal, synchronic, epistemic coherence requirements (of ideal rationality).

[25] This includes Kaplan's (2013) line on the Preface, which appeals to the norms of "what we are willing to say in the context of inquiry". According to Kaplan, "what we are willing to say in the context of inquiry" is governed by a requirement of deductive cogency, which is stronger than (CB). Cogency implies (CB) plus closure (under competent deduction). John (the protagonist of our Homogeneous Preface Paradox) does not seem to us to be violating any norms of "what we are willing to say in the context of inquiry". Indeed, it seems to us that nothing prevents John from being a perfectly rational scientific inquirer. In the Negative Phase, we will return to the question of whether there are some contexts (*e.g.*, certain sorts of "contexts of inquiry") in which deductive consistency (or even full cogency) should be enforced.

[26] Other authors besides Christensen (2004), Kolodny (2007), Foley (1992) and Klein (1985) have claimed that paradoxes of consistency place pressure on the claim that (EB) entails (CB). For instance, Kyburg (1970) maintains that the lottery paradox supports (†). We are focusing on preface cases here, since we think they are more compelling than lottery cases (see *fn.* 38).

[27] Strictly speaking, Christensen (2004) never explicitly endorses ($\star$) or (C_2). He focuses on deductive consistency as a coherence-explanans, and he argues that *it* can be "eliminated" from such explanations, in favor of appeals only to probabilistic credal constraints. Having said that, our "straw man Christensen" makes for a clearer and more illuminating contrast.

(C_2) Epistemic phenomena that appear to be adequately explainable only by appeal to coherence requirements for belief (and facts about an agent's beliefs) can be adequately explained entirely by appeal to probabilism (and facts about an agent's credences).

We agree with Christensen about (C_1). In fact, our framework for grounding coherence requirements for belief is inspired by analogous arguments for probabilism as a coherence requirement for partial belief. We will return to this important parallel below. Christensen's (C_2) is part of an error theory regarding epistemological explanations that appear to involve coherence requirements for belief as (essential) explanans. Some such error theory is needed — given ($\star$) — since epistemologists often seem to make essential use of such coherence-explanans.

Kolodny (2007), on the other hand, accepts the following pair:

(K_1) No attitudes (belief, partial belief, or otherwise) are subject to any formal, synchronic, epistemic coherence requirements (of ideal rationality).

(K_2) Epistemic phenomena that appear to be adequately explainable only by appeal to coherence requirements for belief (together with facts about an agent's beliefs) can be adequately explained entirely by appeal to the Evidential Norm (EB), together with facts about an agent's beliefs.

Kolodny's (K_1) is far more radical than anything Christensen accepts. Of course, (K_1) entails ($\star$), but it also entails universal eliminativism about coherence requirements in epistemology. Kolodny doesn't think there are any (ineliminable) coherence requirements (or any ineliminable requirements of ideal rationality, for that matter), period. He doesn't even recognize probabilism as a coherence requirement for credences. As a result, Kolodny needs a different error theory to "explain away" the various epistemological explanations that seem to appeal essentially to coherence requirements for belief. His error theory [(K_2)] uses the Evidential Norm for belief (EB), along with facts about the agent's beliefs, to explain away such appeals to "coherence requirements". So, Kolodny's error theory differs from Christensen's in a crucial respect: Kolodny appeals to local/narrow-scope norms for belief to explain away apparent uses of coherence requirements for belief; whereas, Christensen appeals to global/wide-scope requirements of partial belief to explain away apparent uses of coherence requirements for belief. This is (partly) because Kolodny is committed to the following general claim:

(K_3) beliefs are an essential (and ineliminable) part of epistemology (*i.e.*, the belief concept is ineliminable from some epistemological explanations).

We agree with Kolodny about (K_3). We, too, think that belief is a crucial (and ineliminable) epistemological concept. (Indeed, this is one of the reasons we are offering a new framework for grounding coherence requirements for belief!) Christensen, on the other hand (at least on our reading, see *fn.* 27), seems to be unsympathetic to (K_3).

One last epistemological principle will be useful for the purposes of comparing and contrasting our views with the views of Christensen and Kolodny.[28]

[28]The "If ..., then ..." in (‡) is a material conditional. That is, (‡) asserts: either ($\star$) or (CB).

(‡) If there are any coherence requirements for belief, then deductive consistency [(CB)] is one of them. [*i.e.*, If $\neg(\star)$, then (CB).]

Christensen and Kolodny both accept (‡), albeit in a trivial way. They both reject the antecedent of (‡) [*i.e.*, they both accept ($\star$)]. We, on the other hand, aim to provide a principled way of rejecting (‡). That is to say, we aim to ground new coherence requirements for belief, which are distinct from deductive consistency. We think this is the proper response to the paradoxes of consistency [and (†)].

In the next section, we will present our formal framework for grounding coherence requirements for (opinionated) belief. But, first, we propose a desideratum for such coherence requirements, inspired by the considerations adduced so far.

($\mathcal{D}$) Coherence requirements for (opinionated) belief should never come into conflict with either alethic or evidential norms for (opinionated) belief. Furthermore, coherence requirements for (opinionated) belief should be entailed by both the Truth Norm (TB) and the Evidential Norm (EB).

In light of (†), deductive consistency [(CB)] violates desideratum ($\mathcal{D}$). If a coherence requirement satisfies desideratum ($\mathcal{D}$), we will say that it is conflict-proof. Next, we explain how to ground conflict-proof coherence requirements for (opinionated) belief.

5. Our (Naïve) Framework and (Some of) its Coherence Requirements

As it happens, our preferred alternative(s) to (CB) were not initially motivated by thinking about paradoxes of consistency. They were inspired by some recent arguments for probabilism as a (synchronic, epistemic) coherence requirement for credences. James Joyce (1998; 2009) has offered arguments for probabilism that are rooted in considerations of accuracy (*i.e.*, in alethic considerations). We will be discussing Joyce's argument(s) in detail in Part III. For now, we will present a general framework for grounding coherence requirements for sets of judgments of various types, including both credences and beliefs. Our unified framework constitutes a generalization of Joyce's argument for probabilism. Moreover, when our approach is applied to belief, it yields coherence requirements that are superior to (CB), in light of preface cases (and other similar paradoxes of consistency).

Applying our framework to finite judgment sets $\mathbf{J} = \{j_1, \ldots, j_n\}$ of type $\mathfrak{J}$ only requires completing three steps. The three steps are as follows.

Step 1. Identify a precise sense in which individual judgments j of type $\mathfrak{J}$ can be (qualitatively) *inaccurate* (or *alethically defective*) at a possible world w.[29]

Step 2. Define a (point-wise) *inaccuracy score* $\mathfrak{i}(j, w)$ for individual judgments j of type $\mathfrak{J}$. This score is a numerical measure of *how inaccurate* (in the sense of Step 1) j is (at w). Then, for each judgment set $\mathbf{J}$ define its *total inaccuracy* at w as the *sum* of the point-wise scores of its members $\mathcal{I}(\mathbf{J}, w) \overset{\text{def}}{=} \sum_i \mathfrak{i}(j_i, w)$.[30]

[29] There may be multiple senses in which individual judgments j of type $\mathfrak{J}$ can be said to be (qualitatively) inaccurate (or alethically defective) at a possible world w. Each of these senses could be used to generate a corresponding formal coherence requirement *via* our three step process. In the case of comparative confidence (Part II), this possible ambiguity in the qualitative notion of inaccuracy will play an important role (more important than in Parts I and III).

[30] We will be discussing a substantive desideratum for inaccuracy measures $\mathfrak{i}$ (called *evidential propriety*) below. But, a basic desideratum for such measures is that they be *truth-directed* in the

> **Step 3**. Adopt a fundamental epistemic principle, which uses $\mathcal{I}(\mathbf{J}, w)$ to ground a (synchronic, epistemic) coherence requirement for judgment sets $\mathbf{J}$ of type $\mathfrak{J}$.

This is all very abstract. To make things more concrete, let's look at the simplest application of our framework — to the case of (opinionated) belief. Recall that our agents will be forming (opinionated) judgments on some salient agenda $\mathcal{A}$, which is a (possibly proper) subset of some finite boolean algebra of propositions. That is, for each $p \in \mathcal{A}$, S either believes p or S disbelieves p, and not both.[31] In this way, an agent can be represented by her *belief set* $\mathbf{B}$, which is just the set of her beliefs (B) and disbeliefs (D) over some salient agenda $\mathcal{A}$. Because we're modeling propositions as sets of (classical) possible worlds, a proposition is true at any world that it contains, and false at any world it doesn't contain. With our (naïve) setup in place, we're ready for the three steps.

Step 1 is straightforward. Here is one precise sense in which an individual belief (or disbelief) can be (qualitatively) inaccurate at a possible world w.

$B(p)$ [$D(p)$] is inaccurate at w just in case p is false [true] at w.

Given the accuracy conditions for B/D, this is the most natural sense of "inaccuracy" to select in Step 1.

Step 2 is also rather straightforward. It seems clear that the two kinds of inaccuracies in this context (*viz.*, false beliefs and true disbeliefs) should receive the same (point-wise) inaccuracy *score*. We will (simply as a matter of convention) assign a score of 1 to each of these two types of inaccuracies, and 0 to judgments which are not inaccurate in this sense. That is, we will adopt the following point-wise inaccuracy score for individual beliefs and disbeliefs.

$$\mathfrak{i}(B(p), w) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } p \text{ is false at } w \\ 0 & \text{otherwise} \end{cases}$$

$$\mathfrak{i}(D(p), w) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } p \text{ is true at } w \\ 0 & \text{otherwise} \end{cases}$$

---

sense of (Joyce, 2009). For our purposes, a point-wise measure of the doxastic inaccuracy $\mathfrak{i}(j, w)$ of an individual judgment j at a world w is truth-directed just in case it assigns an inaccuracy of zero if j is (perfectly) accurate (at w), and an inaccuracy greater than zero if j is inaccurate (at w).

[31] Our assumption of opinionation, relative to a salient agenda $\mathcal{A}$, results in no significant loss of generality for present purposes. As we have explained above, we do not think suspension of belief (on the Preface agenda — there are many propositions outside this agenda on which it may be reasonable to suspend) is an evidentially plausible way of responding to the Preface Paradox. Consequently, one of our present aims is to provide a response to paradoxes of consistency that allows for full opinionation (on the salient agendas). Moreover, there are other applications of the present framework for which opinionation is required. Briggs et al. (2014) show how to apply the present framework to the paradoxes of judgment aggregation, which presuppose opinionation on the salient agendas. Finally, the naïve framework we present here can be generalized in various ways. Specifically, generalizing the present framework to allow for suspension of judgment (on the salient agendas) is, of course, desirable (Sturgeon, 2008; Friedman, 2013). See (Easwaran 2013b) for a generalization of the present framework which allows for suspension of judgment on the salient agendas (see *fn.* 39). We'll return to Easwaran's models of suspension in the Negative Phase below.

Then, we will define the *total inaccuracy* of a belief set $\mathbf{B} = \{b_1, \ldots, b_n\}$ (at w) as the sum of the of the point-wise inaccuracy scores of its members. That is:

$$\mathcal{I}(\mathbf{B}, w) \overset{\text{def}}{=} \sum_i \mathfrak{i}(b_i, w)$$

To be sure, this is a very naïve measure of total inaccuracy. It just *counts the number of inaccurate judgments* in $\mathbf{B}$ at w.[32] In this sense, our inaccuracy measure assigns "equal weight" (or "equal importance") to each of the propositions in the agenda $\mathcal{A}$. As it turns out, however, our arguments will not depend on this "equal weight" assumption. We will return to this issue below (and in the Negative Phase).

Step 3 is the philosophically most important step. Before we get to our favored fundamental epistemic principle(s), we will digress briefly to discuss a stronger fundamental epistemic principle that one might find (*prima facie*) plausible. Given our naïve setup, it turns out that there is *a* choice of fundamental epistemic principle that yields deductive consistency [(CB)] as a coherence requirement for opinionated belief. Specifically, consider the following principle:

> **Possible Vindication** (PV). There exists some possible world w at which none of the judgments in $\mathbf{B}$ are inaccurate. Or, to put this more formally, in terms of our inaccuracy measure $\mathcal{I}$: $(\exists w)[\mathcal{I}(\mathbf{B}, w) = 0]$.

Given our naïve setup, it is easy to show that (PV) is equivalent to (CB).[33] As such, a defender of (CB) would presumably find (PV) attractive as a fundamental epistemic principle. However, as we have seen in previous sections, preface cases (and other paradoxes of consistency) have led many philosophers (including us) to reject (CB) as a rational requirement. This motivates the adoption of fundamental principles that are weaker than (PV). Interestingly, as we mentioned above, our rejection of (PV) was not (initially) motivated by Prefaces and the like. Rather, our adoption of fundamental principles weaker than (PV) was motivated (initially) by analogy with Joyce's argument(s) for probabilism as a coherence requirement for credences.

In the case of credences, the analogue of (PV) is *clearly* too strong. In that context (as we'll see in Part III), possible vindication would require that an agent assign *extremal* credences to *all* propositions. One doesn't need Preface cases (or any other subtle or paradoxical cases) to see that this would be an unreasonably strong (rational) requirement. It is for this reason that Joyce (and all others who argue in this way for probabilism) back away from the analogue of (PV) to strictly weaker epistemic principles — specifically, to accuracy-dominance avoidance principles, which are credal analogues of the following fundamental epistemic principle.

---

[32] This is the *Hamming distance* (Deza & Deza 2009) between $\mathbf{B}$ and the set $\mathring{\mathbf{B}}_w$ of accurate judgments (on agenda $\mathcal{A}$) at w. The set $\mathring{\mathbf{B}}_w$ is sometimes called the *vindicated* judgment set (on $\mathcal{A}$ at w). In this context, there is always a *unique* vindicated judgment set. However, as we'll see in Part II, there can be applications of our framework in which this uniqueness condition fails.

[33] Here, we're assuming a slight generalization of the standard notion of consistency. Standardly, consistency applies only to beliefs (not disbeliefs), and it requires that there be a possible world in which all the agent's beliefs are true. More generally, we may define consistency as the existence of a possible world in which all the agent's judgments (both beliefs and disbeliefs) are accurate. Given this more general notion of consistency, (PV) and (CB) are equivalent.

Weak Accuracy-Dominance Avoidance (WADA). $\mathbf{B}$ is not weakly [34] dominated in accuracy. Or, to put this more formally (in terms of $\mathcal{I}$), there does not exist an alternative belief set $\mathbf{B}'$ such that:

(i) $(\forall w)[\mathcal{I}(\mathbf{B}', w) \leq \mathcal{I}(\mathbf{B}, w)]$, and

(ii) $(\exists w)[\mathcal{I}(\mathbf{B}', w) < \mathcal{I}(\mathbf{B}, w)]$.

(WADA) is a very natural principle to adopt, if one is not going to insist that — as a requirement of rationality — it must be possible for an agent to achieve perfect accuracy in her doxastic state. In the credal case, the analogous requirement was clearly too strong to count as a rational requirement. In the case of belief, one needs to think about Preface cases (and the like) to see why (PV) is too strong. Retreating from (PV) to (WADA) is analogous to what one does in decision theory, when one backs off a principle of maximizing (actual) utility to some less demanding requirement of rationality (*e.g.,* dominance avoidance, maximization of expected utility, minimax, *etc.*).[35] Of course, there is a sense in which "the best action" is the one that maximizes actual utility; but, surely, maximization of actual utility is not a rational requirement. Similarly, there is clearly a sense in which "the best doxastic state" is the perfectly accurate [(TB)], or possibly perfectly accurate [(CB)/(PV)], doxastic state. But, in light of the paradoxes of consistency, (TB) and (CB) turn out not to be rational requirements either. One of the main problems with the existing literature on the paradoxes of consistency is that no principled alternative(s) to deductive consistency have been offered as coherence requirements for belief. Such alternatives are just what our Joyce-style arguments provide.

If a belief set $\mathbf{B}$ satisfies (WADA), then we say $\mathbf{B}$ is non-dominated. This leads to the following, new coherence requirement for (opinionated) belief

(NDB) All (opinionated) agents S should (at any given time t) have sets of beliefs (and disbeliefs) that are non-dominated.

Interestingly, (NDB) is strictly weaker than (CB). Moreover, (NDB) is weaker than (CB) in an appropriate way, in light of our Preface Paradoxes (and other similar paradoxes of consistency). Our first two theorems (each with an accompanying definition) help to explain why.

The first theorem states a necessary and sufficient condition for (*i.e.*, a characterization of) non-dominance: we call it *Negative* because it identifies certain objects, the non-existence of which is necessary and sufficient for non-dominance. The second theorem states a sufficient condition for non-dominance: we call it

[34] Strictly speaking, Joyce opts for the apparently weaker principle of avoiding *strict* accuracy dominance. However, in the credal case (assuming continuous, *strictly* proper scoring rules, which is what most people actually use in this context), there is no difference between weak and strict dominance (Schervish *et al.* 2009). In this sense, there is no serious disanalogy (*fn.* 101). Nonetheless, it is worth noting that, in the case of belief (and comparative confidence), there is a significant difference between weak and strict dominance. This difference will be discussed in some detail in §6 below. In general, when we say "dominated" what we mean is *weakly* dominated in the sense of (WADA).

[35] The analogy to decision theory could be made even tighter. We could say that being accuracy-dominated reveals that you are in a position to recognize *a priori* that another option is guaranteed to do better at achieving the "epistemic aim" of getting as close to the truth as possible (compatibly with one's *evidential* requirements). This decision-theoretic stance dovetails nicely with the sentiments expressed by Foley (*op. cit.*). See §8 and the Negative Phase for further discussion.

Positive because it states that in order to show that a certain belief set **B** is non-dominated, it's enough to construct a certain type of object.

DEFINITION 1.1 (Witnessing Sets). **S** *is a* ***witnessing set*** *iff (a) at every world w, at least half of the judgments*[36] *in* **S** *are inaccurate; and, (b) at some world more than half of the judgments in* **S** *are inaccurate.*

THEOREM 1.1 (Negative). **B** *is non-dominated iff* **B** *contains no witnessing set.* [*We will use "(NWS)" to abbreviate the claim that "no subset of* **B** *is a witnessing set." Thus, Theorem 1 can be stated equivalently as:* **B** *is non-dominated iff (NWS).*]

It is an immediate corollary of this first theorem that if **B** is deductively consistent [*i.e*, if **B** satisfies (PV)], then **B** is non-dominated. After all, if **B** is deductively consistent, then there is a world w such that no judgments in **B** are inaccurate at w (*fn.* 33). However, while deductive consistency guarantees non-dominance, the converse is not the case, *i.e.*, non-dominance does not ensure deductive consistency. This will be most perspicuous as a consequence of our second theorem.

DEFINITION 1.2. *A probability function* Pr **represents** *a belief set* **B** *iff for every* $p \in \mathcal{A}$

(i) **B** *contains* $B(p)$ *iff* $\Pr(p) > 1/2$, *and*

(ii) **B** *contains* $D(p)$ *iff* $\Pr(p) < 1/2$.

THEOREM 1.2 (Positive). **B** *is non-dominated if there exists a probability function* Pr *that represents* **B**.[37]

To appreciate the significance of Theorem 1.2, it helps to think about a standard lottery case.[38] Consider the example of the fair lottery that we described in the introduction (it has $n \geq 3$ tickets, exactly one of which is the winner). For each $j \leqslant n$, let p_j be the proposition that the j^{th} ticket is not the winning ticket; let q be the proposition that some ticket is the winner; and, let these $n + 1$ propositions exhaust the agenda $\mathcal{A}$. (Note that the agenda leaves out conjunctions of these propositions.) Finally, let LOTTERY be the following opinionated belief set on $\mathcal{A}$:

$$\text{LOTTERY} \stackrel{\text{def}}{=} \left\{B(p_j) \mid 1 \leqslant j \leqslant n\right\} \cup \{B(q)\}.$$

[36]Here, we rely on naïve counting. This is unproblematic, since all of our agendas and algebras are finite. Moreover, the coherence norm we'll propose in the end (see §7) will not rely on naïve counting and (as a result) will be applicable to both finite and infinite belief sets. All Theorems from Part I are proved in APPENDIX B.

[37]The question: "Does a belief set **B** have a representing probability function?" is decidable (Fitelson 2008). So is the question "Does a belief set **B** have a witnessing set?". This suffices to ensure that coherence properties (in our sense) of (finite) belief sets are *formal*.

[38] We are not endorsing the belief set LOTTERY in this example as epistemically rational. Indeed, we think that the lottery paradox is not as compelling — as a counterexample to (EB) ⇒ (CB) — as the preface paradox is. On this score, we agree with Pollock (1990) and Nelkin (2000). We are just using LOTTERY to clarify the logical relationship between (CB) and (NDB).

In light of Theorem 1.2, LOTTERY is non-dominated. The probability function that assigns each ticket equal probability of winning represents LOTTERY. However, LOTTERY is not deductively consistent. Hence, (NDB) is strictly weaker than (CB).[39]

Not only is (NDB) weaker than (CB), it is weaker than (CB) in a desirable way. More precisely, in accordance with desideratum ($\mathcal{D}$), we will now demonstrate that (NDB) is entailed by both alethic considerations [(TB)/(CB)] and evidential considerations [(EB)]. While there is considerable disagreement about the precise content of the Evidential Norm for belief (EB), there is widespread agreement (at least, among "probabilist evidentialists") that the following is a necessary condition for satisfying (EB).

> **Necessary Condition for Satisfying (EB).** **B** satisfies (EB), *i.e.*, all judgments in **B** are supported by the total evidence, ***only if***:
>
> ($\mathcal{R}$) There exists some probability function that probabilifies (*i.e.*, assigns probability greater than $1/2$ to) each belief in **B** and dis-probabilifies (*i.e.*, assigns probability less than $1/2$ to) each disbelief in **B**.

Most evidentialists agree that probabilification — relative to some probability function — is a minimal necessary condition for justification. This is because most evidentialists accept some precisification of the slogan "probabilities reflect evidence" (*i.e.*, they are what we will call *probabilistic evidentialists*). Admittedly, there is plenty of disagreement about which probability function is implicated in ($\mathcal{R}$).[40] But, because our Theorem 1.2 only requires the existence of *some* probability function that probabilifies S's beliefs and dis-probabilifies S's disbeliefs, it is sufficient to ensure (on most evidentialist views) that (EB) entails ($\mathcal{R}$). Assuming we're correct in our assessment that Prefaces (and other similar paradoxes of consistency) imply (†), this is precisely the entailment that fails for (CB), and the reason why (CB) fails to satisfy desideratum ($\mathcal{D}$) [*i.e.*, why (CB) fails to be conflict-proof], while (NDB) does satisfy it.[41] Thus, by grounding coherence for beliefs in the same way Joyce grounds probabilism for credences, we are naturally led to coherence requirements for (opinionated) belief that are plausible alternatives to (CB).[42] This gives us a principled way to accept (†) while rejecting (‡), and it paves the way for a novel and

[39] It is worth noting that belief sets like LOTTERY can remain non-dominated even if we allow alternative judgment sets to suspend judgment on some or all of the propositions in the salient agenda (Easwaran, 2013b). This is yet another reason why our restriction to opinionated judgment sets (over appropriate agendas of propositions) results in no significant loss of generality (see *fn.* 31). We will return to this point in the Negative Phase below.

[40] Internalists like Fumerton (1995) require that the function $\Pr(\cdot)$ which undergirds (EB) should be "internally accessible" to the agent (in various ways). Externalists like Williamson (2000) allow for "inaccessible" evidential probabilities. And, subjective Bayesians like Joyce (2005) say that $\Pr(\cdot)$ should reflect the agent's credences. Despite this disagreement, most evidentialists agree that (EB) entails ($\mathcal{R}$), which is all we need for present purposes.

[41] Another way to see why there can't be preface-style counterexamples to (NDB) is to recognize that such cases would have to involve not only the (reasonable) belief that *some* of one's beliefs are false, but (unreasonable) belief that *most* of one's beliefs are false.

[42] We have given a general, theoretical argument to the effect that the Evidential Norm for belief [(EB)] entails the coherence requirement(s) for belief that we favor. We know of no analogous general argument for credences. In (Easwaran & Fitelson 2012), we raised the possibility of counterexamples to the analogous theoretical claim the evidential norm for credences (independently) entails probabilism. Joyce (2013) and Pettigrew (2013a) take steps toward general arguments for this claim. We will return to this disanalogous aspect of Joyce's argument in Part III.

compelling response to the Preface (and other similar paradoxes of consistency). Figure 1.1 depicts the logical relations between the epistemic requirements and norms we have discussed so far.

$$
\begin{tikzcd}
(\mathrm{TB}) \arrow[d, Rightarrow] & & (\mathrm{EB}) \arrow[ddl, Rightarrow] \\
(\mathrm{CB})\,/\,(\mathrm{PV}) \arrow[dr, Rightarrow] & & \\
& (\mathcal{R}) \arrow[d, Rightarrow] & \\
(\mathrm{NWS}) \arrow[r, Leftrightarrow] & (\mathrm{NDB})\,/\,(\mathrm{WADA}) &
\end{tikzcd}
$$

FIGURE 1.1. Logical relations between requirements and norms (so far)

In the next two sections, we will elaborate on the family of coherence requirements generated by our framework.

6. A Family of New Coherence Requirements for belief

The analysis above revealed two coherence requirements that are strictly weaker than deductive consistency: ($\mathcal{R}$) and (NDB). There is, in fact, a large family of such requirements. This family includes requirements that are even weaker than (NDB), as well as requirements that are stronger than ($\mathcal{R}$). Regarding the former, the most interesting requirement that is weaker than (NDB) is generated by replacing weak-accuracy-dominance avoidance with strict-accuracy-dominance avoidance, *i.e.*, by adopting (SADA), rather than (WADA), as the fundamental epistemic principle.

> **Strict Accuracy-Dominance Avoidance** (SADA). **B** is not strictly dominated in accuracy. Or, to put this more formally (in terms of $\mathcal{I}$), there does not exist an alternative belief set $\mathbf{B}'$ such that:
>
> $$(\forall w)[\mathcal{I}(\mathbf{B}', w) < \mathcal{I}(\mathbf{B}, w)].$$

It is obvious that (WADA) entails (SADA). That the converse entailment does not hold can be shown by producing an example of a doxastic state that is weakly, but not strictly, dominated in $\mathcal{I}$-inaccuracy. We present such an example in APPENDIX B. What about requirements "in between" (CB) and ($\mathcal{R}$)?

One way to bring out requirements that are weaker than (CB) but stronger than ($\mathcal{R}$) is to think of (CB) and ($\mathcal{R}$) as "limiting cases" of the following parametric family.

> **Parametric Family of Probabilistic Requirements Between ($\mathcal{R}$) and (CB)**
>
> ($\mathcal{R}_r$) There exists a probability function Pr such that, for every $p \in \mathcal{A}$:
>
> (i) **B** contains $B(p)$ iff $\Pr(p) > r$, and
>
> (ii) **B** contains $D(p)$ iff $\Pr(p) < 1 - r$,

where $r \in [1/2, 1)$.

What we have been calling ($\mathcal{R}$) is (obviously) equivalent to member ($\mathcal{R}_{1/2}$) of the above family. And, as the value of r approaches 1, the corresponding requirement ($\mathcal{R}_r$) approaches (CB) in logical strength. This gives rise to a continuum of coherence requirements that are "in between" (CB) and ($\mathcal{R}$) in terms of their logical strength. (CB) is equivalent to the following, extremal probabilistic requirement.

Extremal Probabilistic Equivalent to (CB)

(CB_{Pr}) There exists a probability function Pr such that, for every $p \in \mathcal{A}$:

(i) **B** contains $B(p)$ iff $\Pr(p) = 1$, and

(ii) **B** contains $D(p)$ iff $\Pr(p) = 0$.

To see that (CB_{Pr}) is equivalent to (CB), note that a belief set **B** is consistent (*i.e.*, possibly perfectly accurate) just in case there is a truth-value assignment function that assigns $\top$ to all p such that $B(p) \in \mathbf{B}$ and $\bot$ to all p such that $D(p) \in \mathbf{B}$. But, this is equivalent to the existence of an indicator function that assigns 1 to all the believed propositions in **B** and 0 to all the disbelieved propositions in **B**. And, such indicator functions just are probability functions of the sort required by (CB_{Pr}).

In the next section, we'll look more closely at our family of new coherence requirements, with an eye toward narrowing the field.

7. A Closer Look at Our Family of New Coherence Requirements

First, we note that there is a clear sense in which (NDB) seems to be too weak. (NDB) doesn't even rule out belief sets that contain contradictory pairs of beliefs. For instance, the belief set $\{B(P), B(\neg P)\}$ on the simple agenda $\{P, \neg P\}$ is not weakly dominated in $\mathcal{I}$-inaccuracy. This can be seen in Table 1.1.

			B		**B**$'$		**B**$''$		**B**$'''$	
	P	$\neg P$	$B(P)$	$B(\neg P)$	$B(P)$	$D(\neg P)$	$D(P)$	$B(\neg P)$	$D(P)$	$D(\neg P)$
w_1	**F**	**T**	$-$	$+$	$-$	$-$	$+$	$+$	$+$	$-$
w_2	**T**	**F**	$+$	$-$	$+$	$+$	$-$	$-$	$-$	$+$

TABLE 1.1. Why (NDB) seems too weak

In Table 1.1, a "+" denotes an accurate judgment (at a world) and a "−" denotes an inaccurate judgment (at a world). As you can see, the belief set $\mathbf{B} \overset{\text{def}}{=} \{B(P), B(\neg P)\}$ contains one accurate judgment and one inaccurate judgment in each of the two salient possible worlds, *i.e.*, $\mathcal{I}(\mathbf{B}, w_1) = 1$ and $\mathcal{I}(\mathbf{B}, w_2) = 1$. None of the other three possible (opinionated) belief sets on $\mathcal{A}$ weakly accuracy-dominates **B**. Specifically, let $\mathbf{B}' \overset{\text{def}}{=} \{B(P), D(\neg P)\}$, $\mathbf{B}'' \overset{\text{def}}{=} \{D(P), B(\neg P)\}$ and $\mathbf{B}''' \overset{\text{def}}{=} \{D(P), D(\neg P)\}$. Then

$$1 = \mathcal{I}(\mathbf{B}, w_1) < \mathcal{I}(\mathbf{B}', w_1) = 2,$$
$$1 = \mathcal{I}(\mathbf{B}, w_2) < \mathcal{I}(\mathbf{B}'', w_2) = 2,$$
$$1 = \mathcal{I}(\mathbf{B}, w_1) = \mathcal{I}(\mathbf{B}''', w_1), \text{ and}$$
$$1 = \mathcal{I}(\mathbf{B}, w_2) = \mathcal{I}(\mathbf{B}''', w_2).$$

Therefore, none of $\mathbf{B}'$, $\mathbf{B}''$ or $\mathbf{B}'''$ weakly accuracy-dominates $\mathbf{B}$, which implies that $\mathbf{B}$ satisfies (NDB). But, intuitively, $\mathbf{B}$ should count as incoherent. After all, $\mathbf{B}$ violates ($\mathcal{R}$), which implies that $\mathbf{B}$ cannot be supported by the total evidence — whatever the total evidence is. This suggests that (NDB) is too weak to serve as "the" (strongest, universally binding) coherence requirement for (opinionated) belief. Indeed, we think a similar argument could be given to show that no requirement that is (strictly) weaker than ($\mathcal{R}$) can be "the" coherence requirement for belief. Dominance requirements like (NDB) have other shortcomings, besides being too weak to serve as the strongest (necessary) requirements of epistemic rationality.

Dominance avoidance conditions like (WADA) and (SADA) are defined in terms of the naïve "mistake-counting" measure of inaccuracy $\mathcal{I}(\mathbf{B}, w)$. Such simple counting measures work fine for finite belief sets, but there seems to be no clear way to apply such naïve distance measures to infinite belief sets. On the other hand, probabilistic requirements like ($\mathcal{R}$) can be applied (in a uniform way) to both finite and infinite belief sets. In Parts II and III we'll return to the problem of infinite agendas (and infinite underlying epistemic possibility spaces), in the contexts of our explications of coherence for comparative confidence and numerical credence.

There is another problem (NDB) inherits from (WADA)'s reliance on the naïve, mistake-counting measure of inaccuracy $\mathcal{I}(\mathbf{B}, w)$. This measure seems to require that each proposition in the agenda receive "equal weight" in the calculation of $\mathbf{B}$'s total inaccuracy. One might (for various reasons) want to be able to assign different "weights" to different propositions when calculating the overall inaccuracy of a doxastic state.[43] An examination of the proof of Theorem 1.2 (see Appendix A) reveals that if a belief set $\mathbf{B}$ satisfies ($\mathcal{R}$), then $\mathbf{B}$ will *minimize expected inaccuracy, relative to its representing (evidential) probability function* Pr. In other words, our naïve mistake counting measure of inaccuracy is *evidentially proper*, where this crucial theoretical property is defined (in general terms) as follows.

DEFINITION 1.3 (Evidential Propriety). *Suppose a judgment set* $\mathbf{J}$ *of type* $\mathfrak{J}$ *is supported by the evidence. That is, suppose there exists some evidential probability function* $\Pr(\cdot)$ *which represents* $\mathbf{J}$ *(in the appropriate sense of "represents" for judgment sets of type* $\mathfrak{J}$*). If this is sufficient to ensure that* $\mathbf{J}$ *minimizes expected inaccuracy (relative to* Pr*), according to the measure of inaccuracy* $\mathfrak{I}(\mathbf{J}, w)$*, then we will say that the measure* $\mathfrak{I}$ *is* ***evidentially proper****.*

This is a very abstract, general definition of evidential propriety (to which we shall return in Parts II and III). In the present case of full belief, the proof of Theorem 1.2 (see Appendix A) establishes that our naïve, mistake counting measure of inaccuracy $\mathcal{I}(\mathbf{B}, w)$ is *evidentially proper*, given our conception of an evidential probabilistic representation of a full belief set [*viz.*, ($\mathcal{R}$)]. This is crucial, since it implies that *no probabilistically representable belief set can be accuracy-dominated.*

Evidential propriety is our most important *desideratum* for epistemic inaccuracy measures. We will make extensive use of evidential propriety throughout this monograph. It is *the* property which ensures that *the alethic and evidential norms both entail accuracy non-dominance.* In other words, evidential propriety is the property which (across all three applications of our framework) guarantees that

[43] Joyce's (1998; 2009) argument(s) for probabilism also allow(s) for different weights to be assigned to different propositions in the calculation of the overall inaccuracy (of a credence function). We'll return to this analogous aspect of Joyce's argument in Part III.

accuracy non-dominance is a formal, wide-scope requirement which is entailed by both alethic and evidential norms (*i.e.*, that accuracy non-dominance is a *conflict-proof* requirement, which satisfies desideratum $\mathcal{D}$).

The proof of our main theorem for full belief (Theorem 1.2) does not rely on the assignment of "equal weights" to each of the propositions in $\mathcal{A}$. In fact, any assignment of (constant) weights to the propositions in $\mathcal{A}$ is compatible with Theorem 1.2. So, another advantage of adopting ($\mathcal{R}$) — as opposed to (NDB) — as "the" coherence requirement for belief is that it allows us to use any additive distance measure we like (*viz.*, to assign any constant weights we like to the propositions in $\mathcal{A}$), while preserving the (Joycean) connection between coherence and the avoidance of accuracy-dominance. In this sense, our adoption of the (Hamming) measure $\mathcal{I}$ (which assigns equal weight to all propositions in $\mathcal{A}$) was merely a conventional choice, since any additive (and truth-directed — *fn.* 30) measure would work.

What about requirements stronger than ($\mathcal{R}$)? For instance, could some requirement ($\mathcal{R}_r$) — with $r > 1/2$ — be a better candidate than ($\mathcal{R}$) for "the" coherence requirement for (opinionated) belief? We suspect this will depend on the context in which a doxastic state is being evaluated. Recall that the key premise in our argument that ($\mathcal{R}$) satisfies desideratum ($\mathcal{D}$) [*i.e.*, that ($\mathcal{R}$) is conflict-proof] was the assumption that the Evidential Norm (EB) entails ($\mathcal{R}$). It is uncontroversial (among probabilist-evidentialists) that this entailment holds in all contexts. What happens to this entailment when we replace ($\mathcal{R}$) with ($\mathcal{R}_r$) and $r > 1/2$? When $r > 1/2$, it will no longer be uncontroversial — even among probabilist-evidentialists — that (EB) entails ($\mathcal{R}_r$) in all contexts. For instance, consider a Lockean, who thinks that one ought to believe P just in case $\Pr(P) > r$, where the value of the threshold r (and perhaps the evidential probability function Pr) depends on context. If $r > 1/2$, then such a probabilist-evidentialist will only accept the entailment (EB) $\Rightarrow$ ($\mathcal{R}_r$) in some contexts. As a result, when $r > 1/2$, the entailment (EB) $\Rightarrow$ ($\mathcal{R}_r$) will not hold uncontroversially, in a context-independent way. However, $r = 1/2$ seems too low to generate a strong enough coherence requirement in most (if not all) contexts, *i.e.*, the requirement ($\mathcal{R}$) = ($\mathcal{R}_{1/2}$) seems too weak in most (if not all) contexts.

To see this, consider the class of minimal inconsistent sets of propositions of size n: $\mathbb{B}_n$. That is, each member of $\mathbb{B}_n$ is an inconsistent set of size n containing no inconsistent proper subset. For instance, each member of $\mathbb{B}_2$ will consist of a contradictory pair of propositions. We've already seen that believing both elements of a member of $\mathbb{B}_2$ is ruled out as incoherent by ($\mathcal{R}$) = ($\mathcal{R}_{1/2}$), but not by (NDB). However, believing each element of a member of $\mathbb{B}_3$ (*e.g.*, believing each of the propositions in $\{P, Q, \neg(P \,\&\, Q)\}$) will *not* be ruled out by ($\mathcal{R}_{1/2}$). In order to rule out inconsistent belief sets of size three, we would need to raise the threshold r to $2/3$. In other words, ($\mathcal{R}_{2/3}$) is the weakest requirement in the ($\mathcal{R}_r$) family that rules out believing each member of a three-element minimal inconsistent set (*i.e.*, each member of some set in $\mathbb{B}_3$). In general, we have the following theorem.[44]

[44]Hawthorne & Bovens (1999) prove some very similar formal results; and Christensen (2004) and Sturgeon (2008) appeal to such formal results in their discussions of the paradoxes of consistency. However, all of these authors presuppose that the probabilities involved in their arguments are the credences of the agent in question. Our probability functions need not be credence functions (see *fn.* 40). Indeed, we need not even assume here that agents have degrees of belief. This is important for us, since we do not want to presuppose any sort of reduction (or elimination) of belief. See the Negative Phase and Parts II and III for further discussion regarding issues of reduction.

THEOREM 1.3. *For all* $n \geq 2$ *and for each set of propositions* $\mathbf{P} \in \mathbb{B}_n$*, if* $r \geq \frac{n-1}{n}$ *then* $(\mathcal{R}_r)$ *rules out believing every member of* $\mathbf{P}$*, while if* $r < \frac{n-1}{n}$*, then* $(\mathcal{R}_r)$ *doesn't rule out believing every member of* $\mathbf{P}$*.*

In preface (or lottery) cases, n is typically quite large. As a result, in order to rule out such large inconsistent belief sets as incoherent, $(\mathcal{R}_r)$ would require a large threshold $r = \frac{n-1}{n}$. For example, ruling out inconsistent belief sets of size 5 *via* $(\mathcal{R}_r)$ requires a threshold of $r = 0.8$, and ruling out inconsistent belief sets of size 10 *via* $(\mathcal{R}_r)$ requires a threshold of $r = 0.9$. We think this goes some way toward explaining why smaller inconsistent belief sets seem "less coherent" than larger inconsistent belief sets.[45] Having said that, we don't think there is a precise "universal threshold" r such that $(\mathcal{R}_r)$ is "the" (strongest universally binding) coherence requirement. There are precise values of r that yield clear-cut cases of universally binding coherence requirements $(\mathcal{R}_r)$, *e.g.*, $r = 1/2$. And, there are precise values of r which yield coherence requirements $(\mathcal{R}_r)$ that we take to be clearly not universally binding, *e.g.*, $r = 1 - \epsilon$, for some minuscule ϵ.[46] What, exactly, happens in between? We'll have to leave that question for a future investigation.[47]

In the next section, we'll elaborate briefly on an illuminating decision-theoretic analogy that we mentioned in passing (*e.g.*, *fn.* 35). Then, we'll tie up a few remaining theoretical loose ends. Finally, we'll consider a troublesome example for our framework, inspired by Caie's (2013) analogous recent counterexample to Joyce's argument for probabilism.

8. A Decision-Theoretic Analogy

If we think of closeness to vindication as a kind of epistemic utility (Pettigrew, 2013b; Greaves, 2013), then we may think of $(\mathcal{R})$ as an expected epistemic utility maximization principle. On this reading, $(\mathcal{R})$ is tantamount to the requirement that an agent's belief set should maximize expected epistemic utility, relative to some evidential probability function. Expected utility maximization principles (even weak ones like $\mathcal{R}$) are stronger (*i.e.*, less permissive) than dominance principles, which (again) explains why $(\mathcal{R})$ is stronger than (NDB).

We could push this decision-theoretic analogy even further. We could think of the decision-theoretic analogue of (TB) as a principle of *actually* maximizing utility (AMU), *i.e.*, choose an act that maximizes utility in the *actual* world. We could

[45] More generally, it seems that the $(\mathcal{R}_r)$ can do a lot of explanatory work. For instance, we mentioned above (*fn.* 7) that even Kolodny — a coherence eliminativist — should be able to benefit from our framework and analysis. We think Kolodny can achieve a more compelling explanatory error theory by taking, *e.g.*, $(\mathcal{R})$, rather than (CB), as his target. Finally, we suspect that debates about the existence of coherence requirements would become more interesting if we stopped arguing about whether (CB) is a requirement and started arguing about $(\mathcal{R})$ or (NDB) [or the other $(\mathcal{R}_r)$] instead.

[46] When we say that $r = 1 - \epsilon$, for some minuscule ϵ, leads to a coherence requirement $(\mathcal{R}_r)$ that is not universally binding, we are not making a claim about any specific probability function (see *fn.* 44). For instance, we're not assuming a Lockean thesis with a "universal" threshold of $r = 1 - \epsilon$. It is important to remember that $(\mathcal{R}_r)$ asserts only the existence of some probability function that assigns a value greater than r to all beliefs in $\mathbf{B}$ and less than $1 - r$ to all disbeliefs in $\mathbf{B}$. This is why $(\mathcal{R}_{1-\epsilon})$ is strictly logically weaker than (CB).

[47] Fully answering this question would require (among other things) a more substantive analysis of the relationship between the Evidential Norm (EB) and the requirements $(\mathcal{R}_r)$, in cases where $r > 1/2$. That sort of analysis is beyond the scope of this monograph.

think of the decision-theoretic analogue of (CB) as a principle of *possibly* maximizing utility (PMU), *i.e.*, choose an act that maximizes utility in some *possible* world. And, as we have already discussed (see *fn.* 35), (NDB)/(WADA) would be analogous to a (weak or pareto) dominance principle in decision theory. The general correspondence between the epistemic norms and requirements discussed above and the analogous decision-theoretic principles is summarized in Table 1.2.[48]

Epistemic Principle	Analogous Decision-Theoretic Principle
(TB)	(AMU) Do ϕ only if ϕ maximizes utility in the *actual* world.
(CB)/(PV)	(PMU) Do ϕ only if ϕ maximizes u in *some possible* world.
($\mathcal{R}$)	(MEU) Do ϕ only if ϕ maximizes *expected* utility.[49]
(NDB)/(WADA)	(WDOM) Do ϕ only if ϕ is *not weakly dominated* in utility.
(SADA)	(SDOM) Do ϕ only if ϕ is *not strictly dominated* in utility.

TABLE 1.2. A Decision-Theoretic Analogy

Like (CB), the principle of "possible maximization of utility" (PMU) is not a requirement of rationality. And, as in the case of (CB), in order to see the failure of (PMU), one needs to consider "paradoxical" cases. See (Parfit 1988; Kolodny & MacFarlane 2010) for discussions of a "paradoxical" decision problem (the so-called "Miner Puzzle") in which the rational act is one that does not maximize utility in any possible world. From this decision-theoretic perspective, it is not surprising that deductive consistency (CB) turns out to be too demanding to be a universally binding rational requirement (Foley, 1992).[50]

9. Tying Up A Few Theoretical Loose Ends

As we have seen, the accuracy-dominance avoidance requirement (NDB) is equivalent to a purely combinatorial condition (NWS) defined in terms of witnessing sets. Similar purely combinatorial conditions exist for some of our other coherence requirements as well. Consider these variations on the concept of a witnessing set.

---

[48]The double line between (CB) and ($\mathcal{R}$) in Table 1.2 is intended to separate rational requirements like ($\mathcal{R}$) from principles like (CB) that are too demanding/substantive to count as universally binding rational requirements. We're not sure exactly how to draw this line, but we think that reflection on how the analogous line can be drawn on the decision-theoretic side of the analogy may shed some light on this question. Specifically, we suspect that the principle of maximization of expected utility (relative to *some* Pr-function — *fn.* 49) seems like a good place to (provisionally) draw this line.

[49]In order to make the analogy tight, we need to precisify (MEU) in the following way:

(MEU*) Do ϕ only if ϕ maximizes expected utility, relative to *some* probability function.

This is weaker (and more formal) than the standard (more substantive) interpretation of (MEU), which involves maximizing expected utility relative to a *specific* probability function (*viz.*, the agent's credence function in the decision situation). But, that is as it should be, since our notion of coherence is meant to be more formal and less substantive than the traditional (MEU) principle.

[50]For a nice survey of some of the recent fruits of this (epistemic) decision-theoretic stance, see (Pettigrew 2013b). And, see (Greaves, 2013; Berker, 2013) for some meta-epistemological worries about taking this sort of epistemic decision-theoretic stance. We will address these worries, as well as related worries raised in (Caie, 2013; Carr, 2013), in the Negative Phase.

DEFINITION 1.4 (Witnessing$_1$ Sets). **S** *is a* ***witnessing$_1$ set*** *iff at every world* w*, more than half of the judgments in* **S** *are inaccurate.*

DEFINITION 1.5 (Witnessing$_2$ Sets). **S** *is a* ***witnessing$_2$ set*** *iff at every world* w*, at least half of the judgments in* **S** *are inaccurate.*

Corresponding to each of these types of witnessing sets is a requirement stating that no subset of **B** should be a witnessing set of that type. To wit:

(NW$_1$S) No subset of **B** is a witnessing$_1$ set.

(NW$_2$S) No subset of **B** is a witnessing$_2$ set.

It turns out that (NW$_1$S) and (NW$_2$S) are intimately connected with (SADA) and ($\mathcal{R}$), respectively. These connections are established by the following two theorems:

THEOREM 1.4. **B** *is non-strictly-dominated iff* **B** *contains no witnessing$_1$ set.* [*In other words, the following equivalence holds: (SADA)* $\Longleftrightarrow$ *(NW$_1$S)* .]

THEOREM 1.5. **B** *is probabilistically representable (in the sense of Definition 2) only if* [51] **B** *contains no witnessing$_2$ set* [*i.e., ($\mathcal{R}$)* $\Longrightarrow$ *(NW$_2$S)*].

This brings us to our penultimate Theorem. Consider the following condition, which requires that there be no contradictory pairs of judgments in a belief set:

(NCP) **B** does not contain any contradictory pairs of judgments. That is, there is no p such that either $\{B(p), B(\neg p)\} \subseteq \mathbf{B}$ or $\{D(p), D(\neg p)\} \subseteq \mathbf{B}$.

THEOREM 1.6. **B** *is probabilistically representable (in the sense of Definition 2) only if* [52] **B** *satisfies both (NDB) and (NCP)* [*i.e., ($\mathcal{R}$)* $\Longrightarrow$ *(NDB & NCP)*].

There are just two more minor theoretical loose ends to tie up. Some philosophers (Williamson, 2000) defend a *Knowledge Norm* for belief.

(KB) All agents S should (at any given time t) have beliefs that are *known*.

Intuitively, the Knowledge Norm entails both the Truth Norm and the Evidential Norm. This explains why it appears at the very top of Figure 1.2, which depicts the known (*fn.* 51) logical relationships between all the requirements and norms discussed in this chapter. There are also two (very weak) requirements that appear at the very bottom of Figure 1.2. In order to explain the relationship between those (final) two requirements, we need one final definition. Let $\mathbf{M}(\mathbf{B}, w)$ denote the set of inaccurate judgments in **B** at world w. We can use $\mathbf{M}(\mathbf{B}, w)$ to characterize the following *bedrock* epistemic requirement.

---

[51] It is an open question whether the converse of Theorem 5 is true [*i.e.*, it is an open question whether ($\mathcal{R}$) $\Longleftarrow$ (NW$_2$S)]. We know there are no small counterexamples to ($\mathcal{R}$) $\Longleftarrow$ (NW$_2$S). If this implication could be established, then it would show that the naïve counting measure of inaccuracy $\mathcal{I}(\mathbf{B}, w)$ is canonical (with respect to the additive measures of inaccuracy). That is, if (NW$_2$S) is equivalent to ($\mathcal{R}$), then a purely combinatorial (naïve, counting) condition is (by Theorem 1.2) sufficient to ensure non-dominance for any additive measure of inaccuracy. That would be a nice result.

[52] Interestingly, the converse is false [*i.e.*, ($\mathcal{R}$) $\not\Leftarrow$ (NDB & NCP)]. See the APPENDIX.

Strong Strict Accuracy-Dominance Avoidance (SSADA). **B** should *not be strongly strictly dominated* in accuracy. Or, more formally (in terms of **M**), there should *not* exist a belief set $\mathbf{B}'$ such that

$$(\forall w)\,[\mathbf{M}(\mathbf{B}', w) \subset \mathbf{M}(\mathbf{B}, w)].$$

(SADA) entails (SSADA), but not conversely. This is because violating (SSADA) entails that there exists a belief set $\mathbf{B}'$ which *not only strictly accuracy dominates* **B**, but *also never makes any mistakes that* **B** *doesn't already make.* Clearly, if S's **B** violates (SSADA), this means S is failing to live up to her epistemic aim of making accurate judgments (no matter how this is aim construed). While (SSADA) is exceedingly weak, it does entail the following norm (see *fn.* 5).

(NCB) **B** contains neither $B(\bot)$ nor $D(\top)$.

In fact, (SSADA) and (NCB) are *equivalent*. And, that's the final theorem of this chapter.

THEOREM 1.7. **B** *satisfies (SSADA) if and only if* **B** *satisfies (NCB).*

That ties up the remaining theoretical loose ends. In the next section (the Negative Phase), we address some potential problems with/shortcomings of our explication of formal coherence for belief.

$$
\begin{tikzcd}
 & & \text{(KB)} \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
 & \text{(TB)} \arrow[d, Rightarrow] & & \text{(EB)} \arrow[ddl, Rightarrow] \\
 & \text{(CB)}/\text{(PV)} \arrow[dr, Rightarrow, "(\mathcal{R}_r)" description] & & \\
\text{(NW}_2\text{S)} \arrow[dd, Rightarrow] & & (\mathcal{R})/(\mathcal{R}_{1/2}) \arrow[ll, Rightarrow] \arrow[d, Rightarrow] & \\
 & & \text{(NDB) \& (NCP)} \arrow[d, Rightarrow] & \\
\text{(NWS)} \arrow[rr, Leftrightarrow] \arrow[d, Rightarrow] & & \text{(NDB)}/\text{(WADA)} \arrow[d, Rightarrow] & \\
\text{(NW}_1\text{S)} \arrow[rr, Leftrightarrow] & & \text{(SADA)} \arrow[d, Rightarrow] & \\
 & & \text{(SSADA)} \arrow[r, Leftrightarrow] & \text{(NCB)}
\end{tikzcd}
$$

FIGURE 1.2. Logical relations between (all) requirements and norms for belief

CHAPTER 2

Belief: Negative Phase

If we don't succeed, we run the risk of failure.

Dan Quayle

1. The Ryan Argument for Consistency

Sharon Ryan (1996) offers an argument for deductive consistency as a (universal) rational requirement.[53] Her argument consists of the following three premises.

> **The Closure of Justified Belief Principle** (CJBP). If S is justified in believing p at t and S is justified at t in believing p entails q, then S is justified in believing q at t.
>
> **The No Known Contradictions Principle** (NKCP). No one is ever justified in believing a statement she knows to be of the form $\ulcorner p \,\&\, \neg p \urcorner$.
>
> **The Conjunction Principle** (CP). If S is justified in believing p at t and S is justified in believing q at t, then S is justified in believing $\ulcorner p \,\&\, q \urcorner$ at t.

Interestingly, the first two premises of Ryan's argument have close analogues in our own framework. (CJBP) is a form of *single-premise closure* of justified belief. (WADA) entails the following claim, which is very similar to (CJBP).

(SPC) Suppose p entails q. Then, no agent's belief set should contain the set $\mathbf{B} = \{B(p), D(q)\}$.

We call this principle (SPC) because it is somewhat similar to *single-premise closure*.[54] To see why (SPC) follows from (WADA), assuming our naïve mistake-counting measure of inaccuracy $\mathcal{I}(\mathbf{B}, w)$, consult Table 2.1. As you can see from Table 2.1, $\mathbf{B}'$ weakly accuracy dominates $\mathbf{B}$ [the second row of the table is ruled-out by the precondition of (SPC)].

Ryan's second premise also has a very close analogue in our framework. Indeed, we have already encountered the following principle.

(NCB) No agent's belief set should contain either $B(\perp)$ or $D(\top)$.

[53] We think Ryan's argument is representative of arguments that have been either explicitly or implicitly endorsed by many authors, including Pollock (1990) and Levi (1967).

[54] Strictly speaking, no closure conditions (or any principles of reasoning or inference or implicit commitment) follow from our framework, since we're concerned only with synchronic epistemic evaluation of sets of explicit commitments. But, we think (SPC) is probably the closest thing to single-premise closure that can be expressed in our framework.

			B		**B′**	
	p	q	$B(p)$	$D(q)$	$D(p)$	$B(q)$
w_1	**T**	**T**	$+$	$-$	$-$	$+$
	T	**F**				
w_2	**F**	**T**	$-$	$-$	$+$	$+$
w_3	**F**	**F**	$-$	$+$	$+$	$-$

TABLE 2.1. Why (SPC) follows from (WADA)

And, we have already seen that (NCB) follows from a fundamental epistemic principle that is *even weaker* than either (WADA) or (SADA) — (SSADA). Therefore, (WADA) entails principles that are very similar to Ryan's first two premises.

It is the third premise of Ryan's argument that has no (grounded) analogue in our framework. One way to see the trouble with (CP) is to note that, when it is combined with (CJBP), it yields the following analogue of *multi-premise closure*.

(MPC) Suppose $\{p_1, \ldots, p_n\}$ entails q. Then no agent's belief set should contain the set $\mathbf{B} = \{B(p_1), \ldots, B(p_n), D(q)\}$.

Of course, (MPC) does *not* follow from any of the (new) coherence requirements we have discussed. In preface cases (or any cases in which an agent's belief set is inconsistent), (MPC) would have the absurd consequence that the agent in question *should not disbelieve any proposition*. This is why we do not accept (CP) as a (universal) epistemic rational requirement.

From our present theoretical perspective, (CP) is best understood as a claim about an agent's *implicit* commitments. According to Ryan (*et. al.*), the problem with agents facing preface cases is that they are implicitly committed to many absurd (dis)beliefs involving conjunctions of claims they are explicitly committed to. In light of preface cases (and perhaps other cases as well), however, we maintain that one cannot simply use deductive logic to determine an agent's implicit commitments (in light of their explicit commitments on a given agenda of propositions). That is, one cannot (in general) simply *deduce* (arbitrary) implicit commitments from a given set of explicit commitments. After all, some deductive consequences of an agent's explicit commitments may fail to be supported by the agent's evidence. More generally, we suspect that there is no (reliable) purely formal/logical technique for determining an agent's implicit commitments, in light of their explicit commitments (Harman, 1986; Steinberger, 2014). However, a proper explication of implicit commitment is beyond the scope of this monograph.

2. Agenda-Relativity & Language Dependence

One of the ways we sidestepped questions regarding suspension of judgment (for more on suspension, see §4 below) was by restricting our evaluations to *agendas* $\mathcal{A}$, which are typically *proper subsets* of some salient (full) Boolean algebra of propositions $\mathcal{B}$ over which the agent is forming judgments. By relativizing our evaluations to agendas on which opinionation seems reasonable, we have been able to avoid (some) questions involving suspension of judgment (but see §4 below). This is a benefit of going "agenda-relative" in our epistemic evaluations. But, agenda-relativity also has some potential costs. One of these potential costs is the possibility that our epistemic evaluations will be *language-dependent*.

Miller (1974) famously argued that Tichý's (1974) definition of verisimilitude (of the predictions of a scientific theory, on an agenda of propositions) is language-dependent. It will be useful to rehearse that dialectic here. In the end, our explication of formal coherence for belief will not be language-dependent (in any threatening sense). But, it will be helpful to see how our measures of inaccuracy (of an agent's judgment set on an agenda) differ from similar measures of verisimilitude (of a set of predictions of a scientific theory on an agenda).

Suppose that the weather outside may be either hot (H) or cold ($\neg H$), either rainy (R) or dry ($\neg R$), and either windy (W) or calm ($\neg W$). And suppose that — in fact — the weather outside is hot, rainy, and windy ($H \,\&\, R \,\&\, W$). Now, let T and T' be competing scientific theories about the weather. T predicts that the weather outside is cold, rainy and windy ($\neg H \,\&\, R \,\&\, W$), while T' predicts that it's cold, dry, and calm ($\neg H \,\&\, \neg R \,\&\, \neg W$). Intuitively, one might think that — while T and T' are both false — T is *closer to the truth* than T', since T gets one of its three predictions wrong ($\neg H$), whereas T' gets all three of three its predictions wrong ($\neg H$, $\neg R$ and $\neg W$). However, it seems that this comparison of verisimilitudes can be *reversed* by re-describing the weather using a different (but expressively equivalent) language.

Suppose we say that the weather is Minnesotan (M) iff it is either hot and rainy or cold and dry (*i.e.*, $M \overset{\text{def}}{=} H \equiv R$), and Arizonan ($A$) iff it is either hot and windy or cold and calm (*i.e.*, $A \overset{\text{def}}{=} H \equiv W$). We may now re-express our theories of the weather, according to whether they predict that the weather is hot, Minnesotan and/or Arizonan. That is, we may replace the three atomic weather propositions $\{H, R, W\}$ with the set $\{H, M, A\}$ and retain exactly the same expressive power. Any hypothesis that can be stated in terms of $\{H, R, W\}$ can be restated in an equivalent form in terms of $\{H, M, A\}$. The truth about the (actual) weather may then be re-expressed as $H \,\&\, M \,\&\, A$. Theory T's predictions may be re-expressed as $\neg H \,\&\, \neg M \,\&\, \neg A$, while theory T''s predictions may be re-expressed as $\neg H \,\&\, M \,\&\, A$. Form this perspective, theory T seems to make three false predictions, while T' seems to make only one false prediction. In this sense, we seem to have reversed the verisimilitude ordering: T' now seems to be closer to the truth than T.

This (Miller-style) argument for the language-dependence of verisimilitude comparisons trades on two key assumptions. First, the argument assumes that scientific theories *predict* (*i.e.*, are implicitly committed to) *whatever* they *entail*. This is a form of *deductive closure* for the predictions of scientific theories. Second, when the predictions of the two competing theories are compared using the HRW-language ($\mathcal{L}_{HRW}$) they are compared with respect to the three propositions $\{H, R, W\}$; but, when the predictions of the two competing theories are compared using the HMA-language ($\mathcal{L}_{HMA}$) they are compared with respect to a *different set* of three propositions $\{H, M, A\} = \{H, H \equiv R, H \equiv W\}$. The first assumption is fair in this context, since we may assume that scientific *theories* are implicitly committed to all of their logical consequences. The second assumption, however, is problematic, since a true *reversal* of verisimilitude assessments should, presumably, involve predictions regarding *one and the same set* of propositions.

We can make this second problem more perspicuous by expanding the original set of predictions, so as to include *all* the salient propositions. That is, the actual weather is such that the following *five* propositions (expressed in $\mathcal{L}_{HRW}$) are all true: $\{H, R, W, H \equiv R, H \equiv W\}$. Theory T gets 2/5 of these predictions right, since it predicts $\{\neg H, R, W, H \not\equiv R, H \not\equiv W\}$. Theory T' also gets 2/5 of these predictions right, since it predicts $\{\neg H, \neg R, \neg W, H \equiv R, H \equiv W\}$. And, these verisimilitude numbers remain the same if we re-describe these sets of five predictions in $\mathcal{L}_{HMA}$. The truth about the weather (expressed in $\mathcal{L}_{HMA}$) is: $\{H, H \equiv M, H \equiv A, M, A\}$. Theory T predicts $\{\neg H, H \equiv M, H \equiv A, \neg M, \neg A\}$, while Theory T' predicts $\{\neg H, H \not\equiv M, H \not\equiv A, M, A\}$. So, no matter which language we use to express this (complete) set of five propositions, the two theories will have exactly the same degree of verisimilitude (*viz.*, 2/5). So, Miller's argument *shifts agendas* of predictions as it shifts languages, and this is how it generates the "reversal".

We can run an analogous, Miller-style "reversal argument" with respect to our (naïve) measure of inaccuracy $\mathcal{I}(\mathbf{B}, w)$. Suppose one agent S's belief set on the agenda $\mathcal{A}_{HRW} = \{H, R, W\}$ is $\mathbf{B}_{HRW} = \{D(H), B(R), B(W)\}$. And, suppose another agent S''s belief set on $\mathcal{A}_{HRW}$ is $\mathbf{B}'_{HRW} = \{D(H), D(R), D(W)\}$. If the *actual* world $w_@$ is such that $H \,\&\, R \,\&\, W$ is true, then S's belief set $\mathbf{B}_{HRW}$ is *less $\mathcal{I}$-inaccurate at $w_@$* than S''s belief set $\mathbf{B}'_{HRW}$, since $\mathbf{B}_{HRW}$ has 1 inaccurate judgment (at $w_@$), while $\mathbf{B}'_{HRW}$ contains 3 inaccurate judgments (at $w_@$).

Now, consider a re-description of S's doxastic state, which uses the language $\mathcal{L}_{HMA}$. We can use $\mathcal{L}_{HMA}$ to re-describe S's doxastic state, as follows: $\mathbf{B}_{HMA} = \{D(H), D(M), D(A)\}$.[55] Similarly, we can re-describe S''s doxastic state in $\mathcal{L}_{HMA}$ as: $\mathbf{B}'_{HMA} = \{D(H), B(M), B(A)\}$. And, $\mathbf{B}_{HMA}$ has 3 (actual) inaccuracies, while $\mathbf{B}'_{HMA}$ has only 1. Thus, using $\mathcal{L}_{HMA}$ rather than $\mathcal{L}_{HRW}$ to describe S's doxastic state would seem to lead to a *reversal* of relative inaccuracies, according to our naïve, mistake-counting measure of inaccuracy $\mathcal{I}$.

This Miller-esque argument rests on the same two assumptions as Miller's original argument. And, in the present context, *both* of Miller's assumptions are inappropriate. As before, there is an illicit *change in agenda*, as we move from $\mathbf{B}_{HRW}$ *vs* $\mathbf{B}'_{HRW}$ to $\mathbf{B}_{HMA}$ *vs* $\mathbf{B}'_{HMA}$. That is, $\mathcal{A}_{HRW} = \{H, R, W\} \neq \{H, H \equiv R, H \equiv W\} = \{H, M, A\} = \mathcal{A}_{HMA}$. And, this already makes the cross-language comparisons inappropriate. Moreover, there is an illicit assumption that (a) S is implicitly committed to $\{D(M), D(A)\}$ in virtue of her explicit commitments to $\{D(H), B(R), B(W)\}$, and (b) S' is implicitly committed to $\{B(M), B(A)\}$ in virtue of her explicit commitments to $\{D(H), D(R), D(W)\}$. Neither of these assumptions is kosher, since our framework crucially does *not* assume anything about the implicit commitments of agents (*e.g.*, that an agent's judgments are *closed* under entailment relations). This is an important difference between scientific theories and their predictions (which, presumably, *are* generally closed under entailment relations) and rational agents and their doxastic commitments (which, as far as we are concerned, are *not* generally closed under entailment relations). We will return to this important difference between theories and (arbitrary) sets of judgments below.

[55] Here, in addition to Miller's two assumptions, we will also have to assume (compatibly with our framework) that our agents are such that $(\forall p)[D(p) \equiv B(\neg p)]$. That is, we will have to assume that our agents disbelieve that p iff they believe that $\neg p$.

3. The "Double Counting" Worry

One might worry that our naïve choice of inaccuracy measure can lead to a kind of *double counting* of differences (or mistakes). Duddy and Piggins (2012, p. 857) articulate the worry in this way (with slight translation into our present notation).

> The most widely used metric in the literature is the Hamming metric. This is simply the number of propositions over which the two individuals disagree. So the distance between $\{B(p), B(q), B(p \,\&\, q)\}$ and $\{B(p), D(q), D(p \,\&\, q)\}$ is 2. But therein lies the problem. The proposition $\neg(p\&q)$ is a logical consequence of p and $\neg q$, and $p\&q$ is a logical consequence of p and q. So, given that the individuals both accept p, the disagreement over $p\&q$ is implied by the disagreement over q. The Hamming metric appears to be double counting because it ignores the fact that the propositions are logically interconnected.

One might think that we have already avoided this objection to our naïve measure of inaccuracy. After all, our preferred coherence requirement ($\mathcal{R}$) allows us to weight propositions *unequally*. So, doesn't that allow us to sidestep this "double counting" worry, by assigning different weights to propositions that are logically related to each other in various ways? Interestingly, no. As it happens, the weightings allowed by ($\mathcal{R}$) would not (in general) be able to take into account all of these relations of relative informativeness between propositions.

In order to properly capture this phenomenon, we would need to have weights that attach not to *individual propositions*, but to *collections of judgments*. Not only would our weighting schemes have to weight $D(q)$ $[B(p)]$ more heavily than $D(p)$ $[B(q)]$ when p $[q]$ is more informative than q $[p]$,[56] but our weighting schemes would also have to be sensitive to which judgments are members of the judgment set under evaluation. And, these sorts of weighting schemes are not among the (constant, proposition-wise) schemes allowed by ($\mathcal{R}$). So, although ($\mathcal{R}$) avoids the "equal weight" problem, it does not avoid this "double counting" problem.

The underlying worry here seems to be that — in addition to accuracy — *informativeness* (both absolute and relative informativeness, with respect to other propositions occurring in the agendas being evaluated) must also be taken into account, when calculating the "epistemic disutility" of a doxastic state. This may be so. But, it does not follow from these sorts of "double counting" intuitions *alone*.

As we have already seen, *some* evidential constraints are *implicitly* taken into account by our accuracy-based requirements. That is to say, *some* sensitivity to (EB) *emerges* from our accuracy-based approach to "epistemic disutility" (without being *explicitly built-in* to the way we measure inaccuracy). Moreover, *some* degree of sensitivity to *informativeness* also emerges from our (naïve) accuracy-based approach. As we explained above, (SPC), which is a close analogue of single-premise closure, follows from accuracy-dominance considerations alone.

We suspect that the objector would like to see *even more* sensitivity to informativeness (*viz*., entailment relations between propositions). But, we do not want to be *so* sensitive to informativeness that we end-up *building-in* a more general (multi-premise) closure condition [*e.g.*, (MPC)], since we think that preface cases show that

[56] More generally, we'd need weighting schemes that involve complicated functions of the logical relations that obtain between various propositions that occur in the judgment sets in question.

general (multi-premise) closure conditions are *not* requirements of rationality. In fairness, we should point out that Duddy and Piggins (2012) are working in the context of judgment aggregation, where it is *assumed* that individual judgment sets are deductively cogent.[57] In this sense, Duddy and Piggins are (like Miller above) operating under presuppositions which beg the central epistemological question of the present monograph. Be that as it may, it remains an interesting open question whether there is a way of defining inaccuracy which yields (MPC) as a consequence of accuracy dominance principles like (WADA) or (SADA).[58]

In the end, the objector may be right that — in addition to accuracy and evidential support — there is another dimension of epistemic evaluation (*viz.*, informativeness) that our framework doesn't adequately capture. If that turns out to be right, then we would need to tweak the framework so that it generates coherence requirements that follow from the Truth Norm, the Evidential Norm *and* the "Informativeness Norm". But, we would need to hear a lot more about this "Informativeness Norm" (if there be such) before we could undertake any such revision of the framework.[59]

4. Suspension of Judgment Revisited

As we have (briefly) explained above, we do not think that our choice to focus on the evaluation of judgment sets over agendas on which opinionation is appropriate leads to a significant loss of generality. This is mainly for two reasons. First, in the kinds of examples that are most important for our dialectical purposes (*e.g.*, prefaces, lotteries, and the like), we think suspension of judgment is simply not warranted with respect to any of the propositions on the salient agendas. Second, even if we were to allow for suspension of judgment, this would not significantly alter the dialectic, since opinionated belief sets would still be rationally permissible (and often epistemically preferable) to judgment sets that include suspensions. In order to appreciate this second reason, we would need to examine a generalization of our framework that relaxes the assumption of opinionation.

Happily, Kenny Easwaran (2013b) has worked-out just such a generalization of the present framework. We won't delve into all of the details of Easwaran's approach. But, it will be worthwhile to examine the basics. The main idea involves defining a more general measure of "epistemic utility," which can be applied to belief, disbelief and suspension of judgment [$S(p)$]. Suspending has a *constant*

---

[57]See Briggs et al. (2014) for an application of the present framework to the aggregation of individual judgment sets that are not necessarily deductively consistent.

[58]We thank Franz Huber, Kit Fine and Bernhard Salow for pressing this question. Another, related open question is whether there exists a definition of inaccuracy which yields (MPC) as a consequence of *minimizing expected inaccuracy*, under some evidential probability function. Perhaps the measure defined in (Duddy and Piggins, 2012) would yield (MPC) as a consequence of accuracy dominance avoidance or minimization of expected inaccuracy [*i.e.*, as a consequence of (WADA) or ($\mathcal{R}$)]. This is an interesting open question. But, because we reject (MPC) as a requirement of rationality, we will not investigate this question further here.

[59]Hempel (1962) uses an "informativeness" measure as his measure of epistemic utility. And, he proves a result [*op. cit.* (12.4)] that is formally very similar to the key lemma in the proof of our central theorem that ($\mathcal{R}$) entails (WADA). See Appendix B. However, like Duddy & Piggins, Hempel and other authors, *e.g.*, Levi (1967), assume deductive cogency as a prior rational constraint on belief sets. Finally, such authors are after explications of narrow-scope norms on belief/acceptance, whereas we are only after explications of formal, wide-scope coherence requirements for belief.

"epistemic utility" (or *score*) of *zero* (the idea being that suspending is neither accurate nor inaccurate), inaccurate opinionated judgments receive a negative score of $-\mathfrak{w}$, and accurate opinionated judgments receive a positive score of $+\mathfrak{r}$, where $\mathfrak{w} \geq \mathfrak{r} > 0$. Then, the *total score* of a judgment set $\mathbf{B}$ (at w) is just the sum of the scores of the individual judgments (at w).

Easwaran proves the following theorem (among many others), which is analogous to our central result above — that $(\mathcal{R})$ entails (WADA).

> **Theorem** (Easwaran, 2013b). An agent S will avoid dominance in total score ***if*** their judgment set $\mathbf{B}$ can be represented as follows:
>
> $(\mathfrak{R})$ There exists a probability function $\Pr(\cdot)$ such that, $\forall p \in \mathcal{A}$:
>
> $$B(p) \text{ iff } \Pr(p) > \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}},$$
> $$D(p) \text{ iff } \Pr(p) < 1 - \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}},$$
> $$S(p) \text{ iff } \Pr(p) \in \left[1 - \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}, \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}\right].$$

Our result that $(\mathcal{R})$ entails (WADA) is (essentially) just a special case of Easwaran's Theorem, where $\mathbf{B}$ contains no suspensions and $\mathfrak{w} = \mathfrak{r} = 1$. In this sense, Easwaran's $(\mathfrak{R})$ is (essentially) just a generalization of our $(\mathcal{R})$.

Armed with Easwaran's generalization, we can see why adding suspension of judgment to the model (in anything like this way) does not significantly alter the way we should think about preface and lottery cases. We'll focus on the lottery case, since it's simpler (but, similar remarks would apply to preface cases). Let's follow the advice of Pollock (and other "*suspensionistas*") and suspend judgment on (some or all of) the lottery propositions. In each possible world w, there will be *exactly one* of our original LOTTERY judgments $B(p_w)$ that is inaccurate. That judgment would contribute $-\mathfrak{w}$ to the total score of LOTTERY. Each of the other opinionated judgments $B(p_{\overline{w}})$ would contribute $+\mathfrak{r}$ to the total score of LOTTERY. If we were to suspend on *all* of the lottery propositions, then the total score of LOTTERY would be $+\mathfrak{r}$ (in every possible world), which would inevitably be *much less* than the total score of our original LOTTERY set (provided only that n is sufficiently large, relative to the ratio between $\mathfrak{w}$ and $\mathfrak{r}$). So, suspending on *all* of the lottery propositions seems to lead to a judgment set LOTTERY′ that is *dominated* in total score by our original, opinionated set LOTTERY. If, on the other hand, we suspend only on *some* of the lottery propositions, then it is possible that we could end-up with a set LOTTERY′ that has a higher total score than LOTTERY (in w). But, this will occur only if we happen to suspend on *exactly the one false lottery proposition in* w (p_w). And, even if we managed to pull of such a "lucky suspension", LOTTERY′ would still only manage to receive a higher total score than LOTTERY *in* w. In the other possible worlds (where the one true lottery proposition is not p_w), LOTTERY′ and LOTTERY may end-up with the same total score (if $\mathfrak{w} = \mathfrak{r}$), but LOTTERY′ can never end-up with a higher total score (since $\mathfrak{w} \geq \mathfrak{r}$).

So, at least on Easwaran's generalization of our framework, suspending judgment on some or all of the lottery propositions does not put us in a better epistemic position (with respect to the wide-scope requirements we've articulated). Perhaps there is another way to model suspension that would confer an epistemic advantage on being non-opinionated in lottery or preface cases, but the onus seems to be on the "*suspensionistas*" to provide one.[60]

5. Contexts in which Consistency *is* a Requirement

We have maintained that deductive consistency is not a universally binding (epistemic) rational requirement. This just means that there are *some* contexts in which inconsistency is (epistemically) rationally *permissible*. Of course, it does *not* follow from this claim that there are *no* contexts in which inconsistency is (epistemically) rationally *im*permissible. Indeed, there seem to be various sorts of contexts in which consistency (or even full cogency) ought to be enforced. We will discuss two such contexts.

The first sort of context in which it is reasonable to enforce consistency (or even full cogency) is what we will call a *demonstrative context*. Suppose S presents what they claim to be a (sound) *demonstration* of some claim q. That is, they announce a set of explicit beliefs $\mathbf{B} = \{B(p_1), \ldots B(p_n)\}$, which constitute the premises of their purported demonstration of q. If S were to learn that some members of $\mathbf{B}$ are inaccurate (*i.e.*, some of the p_i are false), then S would be under pressure to *retract* their purported demonstration. That is, in such a context, S is not only (explicitly) committed to the individual judgments in $\mathbf{B}$, but also (implicitly) to the truth of the *conjunction* of the p_i, since the p_i are meant to be premises in a (sound) *demonstration* of q. In such a context, the explicit judgments in $\mathbf{B}$ have what we will call *argumentative* (or *demonstrative*) *unity*, which undergirds the enforcement of their consistency (or perhaps even their full cogency).

The second sort of context in which consistency is reasonably enforced is what we will call a *theoretical context*. A theoretical context is a context in which some salient belief set $\mathbf{B}$ of an agent has a sufficient degree of *theoretical unity*. A paradigm case would be a context in which an agent is expressing beliefs, the contents of which make-up (or are entailed by) some *theory* they accept. For instance, someone who accepts a scientific theory (*e.g.*, Newton's theory of motion) will generally be committed not only to individual parts of the theory in question, but also anything that is *entailed* (or *predicted*) by that theory (or their explicit commitments regarding parts of the theory). Recall our discussion of Miller's language-dependence argument, above. There, it was crucial that scientific theories are assumed to be "committed" not only to certain (explicit) individual statements that comprise or

[60]To be fair to the "*suspensionistas*", if the penalty for inaccurate judgment $\mathfrak{w}$ is *large enough* (relative to $\mathfrak{r}$), then some judgment sets including suspensions will (inevitably) have greater expected epistemic utility than some inconsistent judgment sets (relative to *any* probability function). For instance, consider the 3-proposition agenda $\mathcal{A} \overset{\text{def}}{=} \{p, q, p \,\&\, q\}$. Provided that $\mathfrak{w} \geq 2 \cdot \mathfrak{r}$, the inconsistent opinionated judgment set $\mathbf{B} \overset{\text{def}}{=} \{B(p), B(q), D(p \,\&\, q)\}$ will *never* maximize expected epistemic utility, relative to *any* probability function; but, the (fully) non-opinionated judgment set $\mathbf{B}' \overset{\text{def}}{=} \{S(p), S(q), S(p \,\&\, q)\}$ will maximize expected epistemic utility, relative to any probability function $\Pr(\cdot)$ such that $\Pr(p)$, $\Pr(q)$ and $\Pr(p \,\&\, q)$ are all on the interval $\left[1 - \frac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}, \frac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}\right]$. However, for sufficiently large judgment sets (like lotteries and prefaces), the requisite discrepancy between $\mathfrak{w}$ and $\mathfrak{r}$ would have to be *very large* in order to secure such a general advantage for *total* suspension (which, owing to the evidential symmetries of the case, would seem to be the only principled way to go). Having said that, it is worth noting that the addition of suspension to our models will significantly affect any verdicts we might reach concerning "the strongest universally binding coherence requirement" among the $(\mathcal{R}_r)$. But, a thorough discussion of these more subtle theoretical effects of adding suspension to our models is beyond our present scope.

characterize the theory (*e.g.*, each of the axioms of the theory), but also to the *consequences* (or predictions) of its "explicit commitments" (or parts).

More generally, in order for it to be reasonable to enforce consistency (or cogency) on a set of judgments, that set must have a certain degree of *unity* (relative to the evaluative context in question). It is useful to think about some contexts in which a judgment set *lacks* the requisite unity. In our Homogeneous Preface case, John's book collects all the (putative) H-relevant evidence that science has come up with. Scientists expect that, in any sufficiently large collection of data (*viz.*, putative evidence), there will be some outliers. In this sense, no scientist expects (or believes) that every claim in every set of data they report (or collect) will be true (Leitgeb, 2014). That is one reason why we chose a *large data set* (as opposed, say, to *parts of a complex scientific theory* that is defended/articulated in a book) as our preface set in this example. To take an even more extreme case, consider the following fanciful example. Suppose we were to select a large subset **B** of some agent S's (rational) beliefs *at random*. Would it be reasonable for us to enforce consistency (or closure) on such arbitrary, randomly sampled sets of S's beliefs? We think not, and we think this is because such randomly sampled beliefs are so *dis*unified. If, on the other hand, we came across a subset of S's beliefs that had the right sort of *unity* (in the context), then it may well be reasonable to enforce consistency (or even cogency) as a requirement on that judgment set (in that context). Collections of beliefs that are *unified* (in the sense we have in mind) can reasonably be taken (in suitable contexts) to represent the world *in concert*.[61]

Finally, it bears repeating that closure/cogency are very crude tools for uncovering the implicit commitments of agents. Just because someone has inconsistent beliefs, that (of course) doesn't mean they are implicitly committed to *everything* (or *nothing*).[62] It is true that if someone were a naïve adherent of closure, then they could "infer anything" once they realized their beliefs were inconsistent. But, as far as we are concerned, this just reveals that facts about the logical consequences of an agent's explicit commitments are not a very useful guide to an agent's implicit commitments (Harman, 1986; Steinberger, 2014). That being said, a proper explication of implicit commitment is beyond the scope of this monograph.

[61] Perhaps (Kaplan, 2013) has some combination of argumentative and theoretical context(s) in mind, when he talks about "the context of inquiry". If so, then we would agree with him to some extent. But, since his view seems to entail the *universal impermissibility* of inconsistency, he must have something much more general in mind (and, to that extent, we disagree).

[62] Pollock (1995) was interested in writing computer programs in `Lisp` that would serve as "artificial reasoners". Because `Lisp` is (essentially) a classical theorem-proving environment (which is based on classical *reductio ad absurdum*), it will *deduce everything* if it's working from an inconsistent database. In this sense, "blind mechanical deduction" *presupposes* classical deductively consistent corpora. I think this is one of the underlying reasons Pollock was so committed to consistency. Various logicians have tried to formulate non-classical logics that allow for (non-trivial) "blind deductive reasoning" from classically inconsistent databases (Belnap, 1977; Priest, 2002). While those alternative logics may be interesting (for some purposes), we think such projects are misguided, since we think it is not a purely formal/logical matter what implicit commitments an agent has (Harman, 1986; Steinberger, 2014). Moreover, we suspect that it also may not be a purely formal/logical matter what predictions are made by an inconsistent scientific theory (Frisch, 2005, pp. 38–39).

6. What About Non-Classical Underlying Logics/Semantics?

As we briefly explained above, we are skeptical that there are good *epistemic* reasons to revise classical logic/semantics (Harman, 1986; Steinberger, 2014). Indeed, the present essay can be viewed as an attempt to provide a response to certain epistemic "paradoxes of consistency" which does *not* (necessarily) involve any revision of classical logic/semantics. However, there may well be good *semantic* reasons to revise classical logic (Fine, 1975; Priest, 2002; Field, 2008; Zardini, 2011). In the case of numerical credence (see Part III), Joyce's argument has been generalized to some non-classical underlying logics/semantics (Williams, 2012). In the case of full belief, analogous generalizations could be constructed.

In order to apply the framework, all we need are suitable explications of (a) the qualitative inaccuracy of an in individual judgment (relative to some suitable point of evaluation) and (b) the point-wise inaccuracy score(s) of the various types of individual judgments (again, relative to some suitable point of evaluation). Once we have (a) and (b) properly explicated, then we can determine the content of the corresponding dominance requirement (WADA). That will yield a new explication of "formal coherence", relative to the non-classical underlying logic/semantics in question. I have not worked out any such generalizations in detail. But, I will briefly discuss one possible generalization along these lines that seems promising.

Let's suppose that a 3-valued logic/semantics is appropriate in some context. Specifically, and simply for the sake of illustration, let's adopt a strong Kleene logic/semantics, according to which there are three truth-values: True, False, and Indeterminate. It would be natural (it seems to me) to suppose that, in such a framework, the analogue of the truth norm would be a norm which required that beliefs be True, disbeliefs be False, and that the "correct/accurate" attitude to have toward Indeterminate claims would be *suspension* of judgment. This gives us a very natural way to define (a) the qualitative inaccuracy of an in individual judgment (presumably, relative to a state of some Kleene algebra), and (b) the point-wise inaccuracy score(s) of the various types of individual judgments (again, relative to some suitable non-classical point of evaluation). Once we have those in hand, we can then impose, *e.g.*, (WADA) as a fundamental epistemic principle, and this will yield some definite formal coherence constraint on judgments in this non-classical setting.[63]

Presumably, similar things could be done for other non-classical underlying logics/semantics — so long as adequate explications of qualitative and quantitative (point-wise) inaccuracy could be given. This is one of many possible generalizations of the present framework.

7. Belief/World Dependence & Epistemic Teleology

We have been (implicitly) presupposing that beliefs have (purely) a mind-to-world direction of fit (Humberstone, 1992). More precisely, we have been (implicitly) assuming that the truth-value of the content of a judgment (being considered

[63] Of course, a full explication of such a "non-classical coherence" requirement would have to involve some generalized notion of "probabilist evidentialism" as well as some generalized notion of "evidential propriety". However, if we could provide such explications, then we could argue that — provided our accuracy score in (b) is *evidentially proper* (in the appropriate, non-classical sense) — the non-dominance requirement we end-up with will be *conflict-proof* (*i.e.*, it will follow from both the alethic and the evidential requirements for the judgments in question).

by an agent S) does not itself depend on whether that judgment is made (by S). However, there are cases in which this sort of "belief/world independence" fails. Some examples of belief/world dependence involve causal dependence (Carr, 2013; Greaves, 2013) while others involve semantic (or conceptual) dependence (Caie, 2013). Because the cases involving semantic/conceptual dependence are simpler (and more perspicuous) in various ways, we will focus on them. Michael Caie (2013) has described such a case in the context of accuracy-based arguments for probabilism as a coherence requirement for numerical credence (which is the subject of Part III of this book). There is a straightforward analogue of Caie's example for the case of full belief. Consider the following (self-referential) claim:

(P) S does not believe that P.

That is, P says of itself that it is not believed by S. Consider the agenda $\mathcal{A} \overset{\text{def}}{=} \{P, \neg P\}$. There seems to be a sound argument for the (worrisome) claim that there are no coherent opinionated belief sets for S on $\mathcal{A}$. This can be seen *via* Table 2.2.

			$\mathbf{B}$		$\mathbf{B}'$		$\mathbf{B}''$		$\mathbf{B}'''$	
	P	$\neg P$	$B(P)$	$B(\neg P)$	$B(P)$	$D(\neg P)$	$D(P)$	$B(\neg P)$	$D(P)$	$D(\neg P)$
w_1	**F**	**T**	−	+	−	−	×	×	×	×
w_2	**T**	**F**	×	×	×	×	−	−	−	+

TABLE 2.2. Caie-style example in our framework

The "×"s in Table 2.2 indicate that these entries in the table are ruled out as logically impossible (given the definition of P). As such it appears that $\mathbf{B}$ and $\mathbf{B}'''$ strictly accuracy-dominate their (live) alternatives (*i.e.*, it appears that $\mathbf{B}$ strictly dominates $\mathbf{B}'$ and $\mathbf{B}'''$ strictly dominates $\mathbf{B}''$). As a result, all of the consistent opinionated belief sets on $\mathcal{A}$ would seem to be ruled out by (SADA). As for $\mathbf{B}$ and $\mathbf{B}'''$, these belief sets consist of contradictory pairs of propositions. We argued earlier that ($\mathcal{R}$) entails (SADA) and rules out any belief set containing contradictory pairs, which seems to mean that ($\mathcal{R}$) rules out all (opinionated) belief sets on $\mathcal{A}$. However, one or both of these entailments might fail in the present case. Our earlier arguments assumed (as standard in probability theory) that none of the propositions exhibit dependence between doxastic states and the truth. Thinking about how to apply ($\mathcal{R}$) in this case requires reconsidering the relationship between a belief set and a probability function.

A probability function assigns numbers to various worlds. When evaluating a belief set with regards to various probability functions, we assumed that worlds specify the truth values of the propositions an agent might believe or disbelieve, while facts about what an agent believes in each world are specified by the belief set, and that every world and every belief set could go together. But with these propositions, there are problems in assessing each belief set with respect to a probability function that assigns positive probability to both worlds, since some belief sets and worlds make opposite specifications of the truth value of a single proposition. For instance, the proposition (P) is specified as false by $\mathbf{B}$ and $\mathbf{B}'$, but as true by w_2, which is why those cells of the table are ruled out.

There are two alternatives that one might pursue. On the first alternative, one considers "restricted worlds" that only specify truth values of propositions that don't concern the beliefs of the agent. On this alternative, there is only one restricted world, and thus only one relevant probability function (which assigns that restricted world probability 1). But this probability function doesn't say anything about the probability of (P), since (P) is a proposition that concerns the beliefs of the agent, and thus is specified by the belief set, and not the restricted worlds and the probability function. So ($\mathcal{R}$) doesn't even apply. But on this alternative, (WADA) and (SADA) still apply, and they rule out $\mathbf{B}'$ and $\mathbf{B}''$, while allowing $\mathbf{B}$ and $\mathbf{B}'''$, even though they involve contradictory pairs.[64]

On the second alternative, one considers full worlds that specify truth values for all propositions, and evaluates each belief set with respect to every world, even though some combinations of belief set and world can't both be actual. On this alternative, Table 2.2 should be replaced by Table 1.1 (*p.* 21). In this new table (which includes impossible pairs), (WADA) and (SADA) no longer rule out $\mathbf{B}'$ and $\mathbf{B}''$. However ($\mathcal{R}$) does rule out $\mathbf{B}$ and $\mathbf{B}'''$. Thus, if one accepts all three of these principles whenever they apply, this second alternative gives the opposite coherence requirement to the first alternative.

This second alternative is endorsed by Briggs (2009, *pp.* 78-83) for talking about actual coherence requirements. She says that an agent with a belief set that is ruled out by the first alternative is "guaranteed to be wrong about something, even though his or her beliefs are perfectly coherent" (*op. cit.*, *p.* 79). The "guarantee" of wrongness only comes about because we allow the specification of the belief set to also specify parts of the world, instead of allowing the two to vary independently.

One might worry that the present context is sufficiently different from the types of cases that Briggs considered that one should go for the first alternative instead. But determining the right interpretation of ($\mathcal{R}$) in the current case is beyond the scope of this book. Caie's examples show that some of our general results may need to be modified in cases where some of the relevant propositions either refer to or depend on the agent's beliefs.[65]

---

[64]Strictly speaking, it is not entirely clear that (WADA) and (SADA) still apply here. Our measure of inaccuracy $\mathcal{I}(\mathbf{B}, w)$ is *evidentially proper* (Theorem 1.2) in ordinary (*i.e.*, non-Caie) cases. But, it seems that this measure *will not be evidentially proper in Caie-style cases.* After all, being probabilistically representable no longer ensures non-dominance (in Caie cases). Of course, the dialectic is a bit tricky here, since Caie is arguing against probabilism as a rational requirement. Be that as it may, it's not entirely clear that it is still appropriate to use the measure $\mathcal{I}(\mathbf{B}, w)$ in these kinds of cases. More generally, it is unclear what to say about how our total evidence bears on claims like (P). There seems to be some kind of *evidential instability* in such cases. We'll return to these subtle aspects of the dialectic in the Negative Phase of Part III.

[65]More generally, these examples involving belief/world dependence raise important questions about the appropriateness of a "teleological" approach to epistemic value/evaluation (Berker, 2013; Greaves, 2013). We are inclined to think that such examples are outside the domain of application the present framework (which involves "purely epistemic" evaluations), since it is concerned with "epistemic acts" that have a *purely mind-to-world* direction of it. This line of defense of the present framework is being pursued by Jason Konek and Ben Levinstein (Konek and Levinstein, 2014). We are sympathetic to their defensive strategy, but a full treatment of these meta-epistemological issues is beyond the scope of the present monograph.

8. What *is* (Full) Belief Anyway?

As we mentioned above, some philosophers would maintain that (full) belief *reduces to* numerical credence (Foley, 2009; Leitgeb, 2013) or comparative confidence (Hawthorne, 2009). Still other philosophers have offered more complex, non-reductionist "pragmatic/decision-theoretic" explications of full belief (Ross and Schroeder, 2012; Buchak, 2013). And, some philosophers would even recommend that we *eliminate* the concept of full belief altogether, perhaps in favor of some more fine-grained attitude (Christensen, 2004; Jeffrey, 1970).

For the purposes of this monograph, we are assuming that belief is not eliminable from epistemology, and that it is subject to its own epistemic norms and requirements (both alethic and evidential). While what we have said so far is not incompatible with certain reductionist views about belief (*e.g.*, certain Lockean theses), it is incompatible with eliminativism about belief (or reductions that imply that consistency is a rational requirement for belief, *etc.*). Presently, we do not wish to take a stand on the question of what full belief *is*. If there are epistemic states that are reasonably subject to the sorts of epistemic evaluations we've been discussing here [*e.g.*, (TB), (EB) and our other normative and evaluative principles], then there will be some epistemological interest in the present models. Finally, we should add that (in general) we favor a *pluralistic* stance, according to which there are various distinct (and irreducible) types of propositional attitudes which are subject to distinct epistemic norms and requirements. In the Negative Phase of Part II we will address some analogous questions about comparative confidence.

Part II

Comparative Confidence[66]

[66] This part of the book is based on joint work with David McCarthy.

CHAPTER 3

Comparative Confidence: Positive Phase

> It has been assumed hitherto as a matter of course that probability is, in the full and literal sense of the word, measurable. I shall have to limit, not extend, the popular doctrine.
>
> John Maynard Keynes

1. Comparative Confidence and its Formal Representation

The contemporary epistemological literature has been focused, primarily, on two types of judgment: full belief and numerical confidence (*viz.*, credence). Not as much attention has been paid to comparative (or relational) epistemic attitudes.[67] In this part of the book, we will be concerned with relational epistemic attitudes that we will call *comparative confidence* judgments. Specifically, we will analyze two kinds of comparative confidence judgments. We will use the notation $\ulcorner p \succ q \urcorner$ to express a comparative confidence judgment which can be glossed as $\ulcorner S$ is *strictly more confident* in the truth of p than they are in the truth of $q\urcorner$, and we will use the notation $\ulcorner p \sim q \urcorner$ to express a comparative confidence judgment which can be glossed as $\ulcorner S$ is *equally confident* in the truth of p and the truth of $q\urcorner$. For the purposes of this investigation, we will make some simplifying assumptions about $\succ$ and $\sim$. These simplifying assumptions are not essential features of our framework, but they will make it easier to develop and explain our justifications of various epistemic coherence requirements for comparative confidence relations.[68]

Our first simplifying assumption is that our agents S form judgments regarding (pairs of) propositions drawn from a finite Boolean algebra of (n) propositions $\mathcal{B}_n$. More precisely, our agents will make comparative confidence judgments regarding (pairs of) propositions on m-proposition *agendas* $\mathcal{A}$, which are (possibly

[67] It wasn't always thus. Keynes (1921), de Finetti (1937, 1951), Koopman (1940), Savage (1972) and Fine (1973) emphasized the importance (and perhaps even fundamentality) of comparative confidence. Moreover, there are some notable recent exceptions to this general trend (Hawthorne, 2009).

[68] It is difficult to articulate the intended meanings of $\ulcorner p \succ q \urcorner$ and $\ulcorner p \sim q \urcorner$ without implicating that these relations reduce to (or essentially involve) some comparison of *non-relational* (and perhaps numerical) credences $b(\cdot)$ of the agent S [*e.g.*, $b(p) > b(q)$ or $b(p) = b(q)$]. But, it is important that no such reductionist assumption be made in the present context. Later in this chapter, we will discuss issues of numerical representability of $\succeq$-relations. And, we will discuss some problems involving numerical reductionism for comparative confidence in the Negative Phase. But, the reader should assume that $\succ$ and $\sim$ are *autonomous* relational attitudes, which may not (ultimately) reduce to (or essentially involve) any non-relational attitude of agents. Other glosses on $\ulcorner p \succ q \urcorner$ ($\ulcorner p \sim q \urcorner$) have been given in the literature, *e.g.*, $\ulcorner S$ judges p to be strictly more believable/plausible than $q\urcorner$ ($\ulcorner S$ judges p and q to be equally believable/plausible$\urcorner$). Another gloss of $\ulcorner p \succ q \urcorner$ ($\ulcorner p \sim q \urcorner$) is $\ulcorner S$'s total evidence strictly favors p over $q\urcorner$ ($\ulcorner S$'s total evidence favors neither p over q nor q over $p\urcorner$).

proper) subsets of $\mathcal{B}_n$ (*viz.*, $m \leq n$). Because we are assuming that the objects of comparative confidence judgments are (pairs of) classical, possible-worlds propositions, we will (as in Part I) be presupposing a kind of (weak) logical omniscience, according to which the the agent is aware of all logical equivalences (and so we may always substitute logical equivalents within comparative confidence judgments).

Next, we adopt some conventions regarding the order-structure of an agent's comparative confidence relation. First, we will assume that the relation $\succ$ constitutes a *strict order* on the agenda in question $\mathcal{A}$. That is, we will assume that $\succ$ satisfies the following two ordering conditions.

> **Irreflexivity of** $\succ$. For all $p \in \mathcal{A}$, $p \not\succ p$.
>
> **Transitivity of** $\succ$. For all $p, q, r \in \mathcal{A}$, *if* $p \succ q$ and $q \succ r$, *then* $p \succ r$.

Second, we assume that $\sim$ is an *equivalence relation* on $\mathcal{A}$. That is, we assume:

> **Reflexivity of** $\sim$. For all $p \in \mathcal{A}$, $p \sim p$.
>
> **Transitivity of** $\sim$. For all $p, q, r \in \mathcal{A}$, *if* $p \sim q$ and $q \sim r$, *then* $p \sim r$.
>
> **Symmetry of** $\sim$. For all $p, q \in \mathcal{A}$, *if* $p \sim q$, *then* $q \sim p$.

Just as in the case of full belief, we will be focusing on agendas $\mathcal{A}$ over which our agents are *opinionated*. More precisely, we will be making the following opinionation assumption regarding our agents' comparative confidence relations.[69]

> **Opinionation of** $\succeq$. For each of the $\binom{m}{2}$ pairs of propositions $p, q \in \mathcal{A}$, our agents will form *exactly one* of the following three possible comparative confidence judgments: *either* $p \succ q$ *or* $q \succ p$ *or* $p \sim q$.

We will use the symbol $\succeq$ to refer to an agent's entire comparative confidence relation over $\mathcal{A}$, *i.e.*, the set of her comparative confidence judgments over $\mathcal{A}$. If $\succeq$ satisfies all five of the ordering assumptions regarding $\succ$ and $\sim$ above, then we will say that $\succeq$ is a *total preorder* on $\mathcal{A}$. We will adopt the notation ⌜$p \succeq q$⌝ to refer to a (generic) individual comparative confidence judgment regarding p and q, which may be either a strict ($\succ$) or an indifference ($\sim$) judgment.

Two final notes are in order here, by way of setup. The combination of (weak) logical equivalence and our ordering assumptions entails the following non-trivial logical omniscience principle, which will play a key role in our proofs below.

(LO) If p and q are logically equivalent, then S judges that $p \sim q$. [And, if S judges $p \succ q$, then p and q are *not* logically equivalent.]

---

[69] We are well aware of the fact that some of these assumptions about the order structure of $\succeq$ (especially, transitivity and opinionation) have been a source of controversy in the literature on coherence requirements for comparative confidence relations. See, for instance, (Forrest, 1989; Fishburn, 1986; Lehrer and Wagner, 1985) for discussion. However, our local, agenda-relative versions of these assumptions are not as controversial as the usual, global versions, which apply to the entire algebra $\mathcal{B}_n$. Having said that, even our agenda-relative ordering assumptions remain somewhat controversial. But, we have chosen (in this Positive Phase) to simplify things by bracketing controversies about the order structure of $\succeq$ on finite agendas $\mathcal{A}$. We will return to some of these questions regarding the order-structure of comparative confidence relations $\succeq$ in the Negative Phase.

Furthermore, we will add the following final background assumption about the kinds of orderings we will be evaluating.

Regularity. For all $p \in \mathcal{A}$, if p is contingent, then $p \succ \bot$ and $\top \succ p$.[70]

Our aim here will be to provide direct, epistemic justifications for various formal coherence requirements — above and beyond the assumption that $\succeq$ is a (Regular) total preorder — that have been proposed for $\succeq$ in the contemporary literature. But, first, we'll need to explain how we're going to formally represent $\succeq$-relations.

One convenient way to represent a $\succeq$-relation on a finite agenda $\mathcal{A}$ containing m propositions is *via* its *adjacency matrix*. Let $p_1, \ldots, p_m$ be the m propositions contained in some agenda $\mathcal{A}$. The adjacency matrix $A^{\succeq}$ of a $\succeq$-relation on $\mathcal{A}$ is an $m \times m$ matrix of 0s and 1s such that $A^{\succeq}_{ij} = 1$ iff $p_i \succeq p_j$.

It's instructive to look at a simple example. Consider the simplest sentential Boolean algebra $\mathcal{B}_4$, which is generated by a single contingent claim P. This algebra $\mathcal{B}_4$ contains the following four propositions: $\langle p_1, p_2, p_3, p_4 \rangle = \langle \top, P, \neg P, \bot \rangle$. To make things concrete, let P be the claim that a coin (which is about to be tossed) will land heads (so, $\neg P$ says that the coin will land tails). Suppose our agent S is equally confident in (*viz.*, epistemically indifferent between) P and $\neg P$. And, suppose that S is strictly more confident in $\top$ than in any of the other propositions in $\mathcal{B}_4$, and that S is strictly less confident in $\bot$ than in any of the other propositions in $\mathcal{B}_4$. This description fully characterizes a (regular, totally preordered) $\succeq$-relation on the agenda consisting of the entire Boolean algebra $\mathcal{B}_4$, which has the adjacency matrix representation (and the graphical representation) depicted in Figure 3.1. In the adjacency matrix $A^{\succeq}$ of $\succeq$, a 1 appears in the $\langle i, j \rangle$-th cell just in case $p_i \succeq p_j$. In the graphical representation of $\succeq$, an arrow is drawn from p_i to p_j only if[71] $p_i \succeq p_j$. With our basic formal framework in hand, we are ready to proceed.

In the next section, we'll discuss a fundamental coherence requirement for $\succeq$ that has been (nearly) universally accepted in the contemporary literature. Then, we will layout our general framework for grounding $\succeq$-coherence requirements, and we will explain how our framework can be used to provide a novel epistemic justification of this fundamental coherence requirement for $\succeq$.

2. The Fundamental Coherence Requirement for $\succeq$

The literature on coherence requirements for $\succeq$ has become rather extensive. A plethora of coherence requirements of varying degrees of strength, *etc.*, have been proposed and defended. We will not attempt to survey all of these requirements here.[72] Instead, we will focus on a particular family of requirements, which can be

[70] In the case of numerical credence, the assumption of (numerical) Regularity (Shimony, 1955) has been a source of significant controversy (Hájek, 2012). However, our present assumption of *comparative* Regularity is significantly less controversial than its numerical analogue (Easwaran, 2014). So, our present assumption of Regularity is not something that most people would find objectionable. We will discuss the relationship between qualitative and quantitative Regularity assumptions in the Negative Phase below. Finally, it is worth noting that in the case of numerical confidence (Part III) no assumption of (numerical) Regularity will be required to apply our framework in a probative way.

[71] We're omitting some arrows in our graphical representations of total preorders. Specifically, because such $\succeq$-relations are transitive, there will be lots more arrows than we're showing in our diagrams. So, we're just using these graphical representations to give the gist of the relation depicted.

[72] See (Halpern, 2003, Ch. 2) for an up-to-date and comprehensive survey. See, also, (Wong et al., 1991; Capotorti and Vantaggi, 2000) and references therein.

$\succeq$	$\top$	P	$\neg P$	$\bot$
$\top$	1	1	1	1
P	0	1	1	1
$\neg P$	0	1	1	1
$\bot$	0	0	0	1

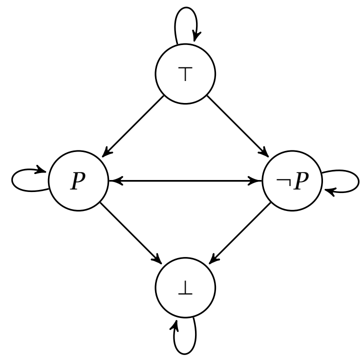

FIGURE 3.1. Adjacency matrix $A^{\succeq}$ and graphical representation of an intuitive $\succeq$-relation on the smallest Boolean algebra $\mathcal{B}_4$

expressed *via* both axiomatic constraints on $\succeq$ and in terms of various kinds of numerical representability of $\succeq$. We begin with the most fundamental of the existing coherence requirements, which is common to almost all the accounts of comparative confidence we have seen. In order to properly introduce this fundamental coherence requirement for comparative confidence, we will first need to introduce the concept of a *plausibility measure* (or a *capacity*) on a Boolean algebra.

A *plausibility measure* (*a.k.a.*, a *capacity*) on a Boolean algebra $\mathcal{B}_n$ is real-valued function $\mathrm{Pl} : \mathcal{B}_n \mapsto [0, 1]$ which maps propositions from $\mathcal{B}_n$ to the unit interval, and which satisfies the following three axioms (Halpern, 2003, *p.* 51).

(Pl_1) $\mathrm{Pl}(\bot) = 0$.

(Pl_2) $\mathrm{Pl}(\top) = 1$.

(Pl_3) For all $p, q \in \mathcal{B}_n$, *if* p entails q, *then* $\mathrm{Pl}(q) \geq \mathrm{Pl}(p)$.

That is, a plausibility measure $\mathrm{Pl}(\cdot)$ is a real-valued function from $\mathcal{B}_n$ to the unit interval, which (a) assigns maximal value to tautologies, (b) assigns minimal value to contradictions, and (c) never assigns logically stronger propositions a greater value than logically weaker propositions. The fundamental coherence requirement for $\succeq$ — which we will call ($\mathfrak{C}$) — can be stated in terms of representability by a plausibility measure. That is, here is one way of stating ($\mathfrak{C}$).

($\mathfrak{C}$) It is a requirement of ideal epistemic rationality that an agent's $\succeq$-relation (assumed to be a regular, total preorder on a finite agenda $\mathcal{A}$) be *representable by some plausibility measure*. That is, a $\succeq$-relation is *coherent only if* there is some plausibility measure Pl such that for all $p, q \in \mathcal{A}$

$$p \succ q \text{ iff } \mathrm{Pl}(p) > \mathrm{Pl}(q), \text{ and } p \sim q \text{ iff } \mathrm{Pl}(p) = \mathrm{Pl}(q).^{73}$$

[73] Strictly speaking, these traditional coherence requirements should be stated *via* axioms (or f-representability constraints) that a relation satisfies *over an entire algebra* $\mathcal{B}_n$. So, strictly speaking, we should be talking about whether a relation $\succeq$ on an agenda $\mathcal{A}$ is *extendible to a relation on the entire algebra* $\mathcal{B}_n$ which satisfies the axioms/representability requirements in question. And, it is extendibility claims such as these that our proofs will establish (or refute). However, for ease of exposition, we will state requirements in terms of constraints on $\succeq$ *over the agenda* $\mathcal{A}$ in question.

It is well known (Capotorti and Vantaggi, 2000) that ($\mathfrak{C}$) can be also stated *via* the following axiomatic constraints on the relation $\succeq$.

($\mathfrak{C}$) It is a requirement of ideal epistemic rationality that an agent's $\succeq$-relation (assumed to be a regular, total preorder on a finite agenda $\mathcal{A}$) satisfy the following two axiomatic constraints

(A_1) $\top \succ \bot$.

(A_2) For all $p, q \in \mathcal{A}$, *if* p entails q, *then* $p \not\succ q$.

In other words, ($\mathfrak{C}$) requires (A_1) that an agent's $\succeq$-relation ranks tautologies *strictly above* contradictions, and (A_2) that an agent's $\succeq$-relation lines up with the "is deductively entailed by" (*viz.*, the "logically follows from") relation.

As far as we know, despite the (nearly) universal acceptance of ($\mathfrak{C}$) as a coherence requirement for $\succeq$, no *epistemic* justification has been given for ($\mathfrak{C}$). Various *pragmatic* justifications of requirements like ($\mathfrak{C}$) have appeared in the literature. For instance, various "Money Pump" arguments and "Representation Theorem" arguments (Savage, 1972; Fishburn, 1986; Halpern, 2003; Icard, 2014) have aimed to show that agents with $\succeq$-relations which *violate* ($\mathfrak{C}$) must exhibit some sort of "pragmatic defect". Once again, following Joyce (1998; 2009), we will be focusing on *non*-pragmatic (*viz.*, *epistemic*) defects implied by the synchronic incoherence of an agent's $\succeq$-relation. Specifically, we will (as in the case of full belief above) be focusing on *alethic* and *evidential* evaluations of an agent's $\succeq$-relation (over a given, finite agenda $\mathcal{A}$), which we take to be *distinctively epistemic*.

In the next section, we'll explain our general (broadly Joycean) strategy for grounding epistemic coherence requirements for $\succeq$. This will allow us to explain *why* ($\mathfrak{C}$) is a requirement of ideal *epistemic* rationality. Moreover, our explanation will be a unified and principled one, which dovetails nicely with similar explications of formal, epistemic coherence requirements for other types of judgment (*e.g.*, as explained in Parts I and III of this book).

3. Grounding the Fundamental Coherence Requirement for $\succeq$

In this section, we show how a natural generalization of Joyce's argument for probabilism can be used to provide a compelling epistemic justification of the fundamental coherence requirement for $\succeq$ [($\mathfrak{C}$)]. This involves going through the Joycean "three steps" (as in §5 of Part I above) — as applied to $\succeq$.

3.1. Step 1: Qualitative inaccuracy of $\succeq$-judgments. We will adopt the (broadly Joycean) idea that a confidence *ordering* is (qualitatively) inaccurate (at w) *iff*[74] it fails to rank all the truths *strictly above* all the falsehoods (at w). Two facts about the inaccuracy (or alethic defectiveness) of *individual* comparative confidence judgments follow immediately from this fundamental assumption about the (qualitative) inaccuracy of comparative confidence orderings.

[74]Joyce would probably not want to accept the only if direction of this biconditional. That is, Joyce would likely want to adopt a weaker notion of qualitative $\succeq$-*in*accuracy, according to which a confidence ordering is inaccurate iff it fails to be identical to the (unique) ordering $\mathring{\succeq}_w$ at w which not only ranks all truths above all falsehoods, but *also ranks propositions with the same truth-value at the same level.* Unfortunately, there is no evidentially proper inaccuracy measure that suits this weaker ($\mathring{\succeq}_w$ non-identity) notion of qualitative $\succeq$-inaccuracy. It is for this reason that we adopt our stronger sense of "inaccurate". We will return to this negative result in the Negative Phase.

Fact 1. If $q \,\&\, \neg p$ is true at w, then $p \succ q$ is inaccurate at w.

Fact 2. If $p \not\equiv q$ is true at w, then $p \sim q$ is inaccurate at w.

The unifying idea behind these two facts is that an individual comparative confidence judgment $p \succeq q$ is *inaccurate/incorrect* at w iff $p \succeq q$ *implies that the ordering* $\succeq$ *of which it is a part is qualitatively inaccurate at* w *(i.e., that some truth fails to be ranked strictly above some falsehood by the relation* $\succeq$ *at* w*).* The two facts above identify the two cases in which this alethic defect is manifest.

Assuming *extensionality*, *i.e.*, that the (qualitative and quantitative) inaccuracy of a comparative confidence ordering over an agenda $\mathcal{A}$ is determined solely by the truth values of the propositions in $\mathcal{A}$ at w, our approach to Step 1 is quite natural (but see *fn.* 74). And, given our particular extensional approach to qualitative $\succeq$-inaccuracy, the two facts above are the *only* (qualitative) facts regarding the (qualitative) inaccuracy of *individual* comparative confidence judgments $p \succeq q$.

Nonetheless, the type of mistake identified in Fact 1 seems *worse* than the type of mistake identified in Fact 2. For the mistake identified in Fact 1 implies that some falsehoods are ranked *strictly above* some truths by the ordering $\succeq$ (in w), whereas the mistake identified in Fact 2 implies only that the truths and falsehoods are *not fully separated* by the ordering $\succeq$ (in w). We will return to this difference between these two kinds of comparative inaccuracies in the next section, when we discuss *quantitative measures of inaccuracy* for $\succeq$-judgments (and sets thereof).

One more remark is in order about the distinction between these two types of comparative inaccuracies. There is something (*prima facie*) peculiar about the type of mistake identified in Fact 2. Recall the relation $\succeq$ depicted in Figure 3.1, above. In that case, we have an agent who is indifferent between the claim that the coin will land heads and the claim that the coin will land tails ($P \sim \neg P$). According to Fact 2, this is *automatically* an *inaccurate* comparative judgment, since $P \equiv \neg P$ is a *logical falsehood.* This may seem somewhat odd, since this relation could certainly be *supported by the agent's evidence* (*e.g.*, if their evidence consisted solely of the claim that the coin is *fair*). What this reveals is that, unlike the case of full belief, in the case of comparative confidence, there can be individual judgments which are supported by some body of evidence — despite the fact that they are (qualitatively) inaccurate in every possible world.[75] We'll return to the distinction between alethic and evidential norms for comparative confidence judgments below. Meanwhile, with our explication of the (qualitative) inaccuracy of individual comparative confidence judgments in hand, we're ready to proceed to Step 2.

3.2. Step 2: Measuring Quantitative $\succeq$-Inaccuracy. Our next step involves the explication of a (point-wise, additive) measure of (total) inaccuracy for comparative confidence relations (or judgment sets). By analogy with the case of full belief, we could opt for a naïve, *mistake-counting* measure of the overall inaccuracy of a comparative confidence relation (*viz.*, judgment set) $\succeq$. That is, we could opt for

$$\Delta(\succeq, w) \stackrel{\text{def}}{=} \text{the number of inaccurate individual judgments in } \succeq \text{ (at } w\text{)}.$$

To understand Δ, it is helpful to recall that a relation $\succeq$ is just a *set* of $\binom{m}{2}$ pairwise comparative confidence judgments, *i.e.*, a set containing *exactly one* of $\{p \succ q, q \succ$

[75]As we'll see in Part III, the same will be true (but *even more starkly*) for numerical credences.

$p, p \sim q\}$ for each of the $\binom{m}{2}$ pairs of propositions p and q in the m-proposition agenda $\mathcal{A}$. In each such set, there will be some number (possibly zero) of inaccurate individual comparative confidence judgments (at w). That number is $\Delta(\succeq, w)$.

We do not think $\Delta(\succeq, w)$ is an appropriate inaccuracy measure (for present purposes). We have two main reservations regarding the use of $\Delta(\succeq, w)$.

- As we explained above, the type of mistake identified in Fact 1 seems *worse* (from an alethic point of view) than the type of mistake identified in Fact 2. But, $\Delta(\succeq, w)$ scores both types of mistakes *equally.*[76]
- Because $\Delta(\succeq, w)$ does not weight the type of mistake identified in Fact 1 *more heavily* than the type of mistake identified in Fact 2, $\Delta(\succeq, w)$ fails to be *evidentially proper* (in the sense of Definition 1.3). Consequently, adopting $\Delta(\succeq, w)$ as our measure of inaccuracy would *run afoul* of the evidential requirements we'll be adopting below.

It is helpful here to work through a couple of simple examples involving $\Delta(\succeq, w)$ in order to (begin to) appreciate these two problems. Recall our toy agent who forms judgments on an agenda $\mathcal{A}$ consisting of the (entire) simplest Boolean algebra $\mathcal{B}_4$. Let S's $\succeq$-relation be given by the ordering depicted in Figure 3.1. There are only two salient possible worlds in this case (w_1 in which P is false, and w_2 in which P is true). It is straightforward to calculate $\Delta(\succeq, w_i)$ in each of these two salient possible worlds. In both worlds, the relation $\succeq$ contains *exactly one* inaccurate indifference judgment: $P \sim \neg P$. And, in both worlds, the relation $\succeq$ does not rank any falsehoods *above* any truths. So, *none* of $\succeq$'s *strict* comparisons is inaccurate in either world. This means that, in both worlds w_i, $\Delta(\succeq, w_i) = 1$. Next, let's consider an *alternative* comparative confidence relation ($\succeq'$) on $\mathcal{B}_4$, as depicted in Figure 3.2. The only difference between $\succeq$ and $\succeq'$ is that $\succeq'$ ranks P

$\succeq'$	$\top$	P	$\neg P$	$\bot$
$\top$	1	1	1	1
P	0	1	1	1
$\neg P$	0	0	1	1
$\bot$	0	0	0	1

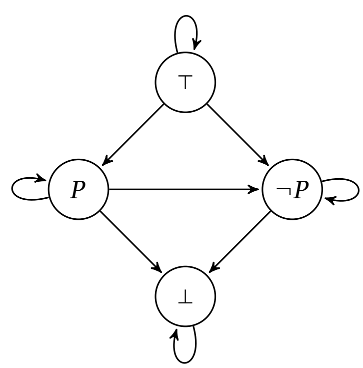

FIGURE 3.2. Adjacency matrix $A^{\succeq'}$ and graphical representation of an alternative relation ($\succeq'$) on the smallest Boolean algebra $\mathcal{B}_4$

[76] This is disanalogous to the case of full belief, where the two types of mistakes (believing falsehoods and disbelieving truths) seem to be *on a par*. And, this explains why we didn't complain about the fact that our point-wise score $\mathfrak{i}$ scored *those* two types of mistakes equally (and it also explains, in part, why our naïve, mistake-counting measure $\mathcal{I}(\mathbf{B}, w)$ was evidentially proper).

strictly above $\neg P$, whereas $\succeq$ is *indifferent* between P and $\neg P$. It is straightforward to compare the values of $\Delta(\succeq, w_i)$ and $\Delta(\succeq', w_i)$ in the two salient possible worlds.

$$\Delta(\succeq, w_1) = 1 = \Delta(\succeq', w_2)$$
$$\Delta(\succeq, w_2) = 1 > 0 = \Delta(\succeq', w_2)$$

As you can see, $\succeq'$ is *never more* inaccurate than $\succeq$, according to Δ; and, $\succeq'$ is *sometimes less inaccurate* than $\succeq$, according to Δ (*i.e.*, in world w_2). In other words, according to Δ, $\succeq'$ *weakly dominates* $\succeq$ in overall inaccuracy. This will mean that (given the initial choice of fundamental epistemic principle that we're going to make in Step 3 below) using Δ would have the effect of *ruling out* $\succeq$ as *epistemically irrational*. And, this is *not a good thing*, since $\succeq$ could (intuitively) be supported by S's evidence (*e.g.*, if her evidence consists of the claim that the coin is *fair*). We will delve further into these (probabilistic) evidential requirements for $\succeq$, below.

Presently, our task is to introduce an alternative measure of $\succeq$-inaccuracy, which avoids both of these problems encountered by Δ. As we will explain later (in the section on evidential requirements for comparative confidence relations), there is (essentially) *only one way* to fix the two problems that plague Δ. Assuming (as a matter of convention) that we assign an inaccuracy score of 1 to the types of ($\sim$) mistakes identified in Fact 2, there is *only one way* to score the types of ($\succ$) mistakes of identified in Fact 1 that leads to an evidentially proper inaccuracy measure, and that is to assign them a score that is *exactly twice as large*. In other words, the following point-wise measure is (essentially) *uniquely* determined by the constraint that the resulting measure of total inaccuracy be evidentially proper.[77]

$$\imath_{\succeq}(p \succeq q, w) \overset{\text{def}}{=} \begin{cases} 2 & \text{if } q \,\&\, \neg p \text{ is true in } w\text{, and } p \succ q, \\ 1 & \text{if } p \not\equiv q \text{ is true in } w\text{, and } p \sim q, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{C} = \{c_1, \ldots, c_{\binom{m}{2}}\}$ be the comparative confidence judgment set determined by a relation $\succeq$ on an m-proposition agenda $\mathcal{A}$. That is, for each pair $p, q \in \mathcal{A}$, $\mathbf{C}$ contains exactly one of the three comparative judgments $p \succ q$, $q \succ p$ or $p \sim q$. In the section on evidential requirements for comparative confidence relations, we will prove that (a) $\imath_{\succeq}(p \succeq q, w)$ undergirds an evidentially proper measure[78]

$$\mathcal{I}_{\succeq}(\succeq, w) \overset{\text{def}}{=} \sum_i \imath_{\succeq}(c_i, w)$$

of the (total) inaccuracy of a relation $\succeq$ (on $\mathcal{A}$), and (b) essentially no other scoring scheme for the two types of mistaken comparative confidence judgments identified

[77] If we were to adopt the stronger notion of "accurate confidence ordering" mentioned in *fn.* 74, then there would be *no* evidentially proper measure of inaccuracy (see the Negative Phase). It is important to note that this result trades essentially on the assumption that our overall inaccuracy measures are *additive*. Konek (2014) is investigating the use of non-additive measures of inaccuracy for comparative confidence judgment sets, in an attempt to apply the framework to the weaker Joycean notion of "inaccuracy" (or non-vindication) of comparative confidence relations.

[78] We will use the notation $\mathcal{I}_{\succeq}(\succeq, w)$ to refer to the total inaccuracy of a relation $\succeq$ on an agenda $\mathcal{A}$. We may also use the notation $\mathcal{I}_{\succeq}(\mathbf{C}, w)$ to refer to the same measure, where it is understood that $\mathbf{C}$ is the set of individual comparative confidence judgments (on $\mathcal{A}$) that comprise the relation $\succeq$.

in Facts 1 and 2 yields an evidentially proper (total) inaccuracy measure for $\succeq$. For now, let's just run with our point-wise and total measures $\mathring{\imath}_{\succeq}$ and $\mathcal{I}_{\succeq}$ above.

3.3. Step 3: The fundamental epistemic principle. We will begin with the same fundamental epistemic principle with which we began in Part I — weak accuracy dominance avoidance.[79]

> **Weak Accuracy-Dominance Avoidance for Comparative Confidence Relations** (WADA$_{\succeq}$). $\succeq$ should *not be weakly dominated* in accuracy. Or, to put this more formally (in terms of $\mathcal{I}_{\succeq}$), there should *not* exist a binary relation[80] $\succeq'$ on $\mathcal{A}$ such that both
>
> (i) $(\forall w)\,[\mathcal{I}_{\succeq}(\succeq', w) \leq \mathcal{I}_{\succeq}(\succeq, q)]$, and
>
> (ii) $(\exists w)\,[\mathcal{I}_{\succeq}(\succeq', w) < \mathcal{I}_{\succeq}(\succeq, w)]$.

Let's return to our toy example(s) above, to illustrate $\mathcal{I}_{\succeq}$-*non*-dominance. Recall our toy agent S who has formed comparative confidence judgments ($\succeq$) on $\mathcal{B}_4$, regarding the toss of a coin (as depicted in Figure 3.1). It can be shown (more on this later) that there is *no* binary relation $\succeq'$ on $\mathcal{B}_4$ that weakly dominates S's $\succeq$ in accuracy (according to $\mathcal{I}_{\succeq}$). We will prove this general claim later on. Meanwhile, we will show that the alternative relation $\succeq'$ (which weakly Δ-dominated $\succeq$) does not weakly $\mathcal{I}_{\succeq}$-dominate $\succeq$. It is easy to see that

$$\mathcal{I}_{\succeq}(\succeq, w_1) = 1 < 2 = \mathcal{I}_{\succeq}(\succeq', w_1)$$
$$\mathcal{I}_{\succeq}(\succeq, w_1) = 1 > 0 = \mathcal{I}_{\succeq}(\succeq', w_2)$$

That is, according to $\mathcal{I}_{\succeq}$, $\succeq$ is more accurate than $\succeq'$ in w_1 but less accurate than $\succeq'$ in w_2. So, neither of these comparative confidence relations weakly dominates the other in accuracy, according to $\mathcal{I}_{\succeq}$.[81]

The avoidance of weak dominance in doxastic inaccuracy is a basic principle of epistemic utility theory. But, as in the case of full belief (see §9 of Part I), some of our results continue to hold, even if we adopt *weaker and more sacrosanct* fundamental epistemic principles than (WADA$_{\succeq}$). As in Part I, we will discuss two requirements that are weaker than (WADA$_{\succeq}$). Predictably, the first is

> **Strict Accuracy-Dominance Avoidance for Comparative Confidence Relations** (SADA$_{\succeq}$). $\succeq$ should *not be strictly dominated* in accuracy. Or, to

[79]In the credal case, weak and strict accuracy dominance avoidance turn out to be (practically) equivalent (see *fn.* 34). As we saw in Part I, in the case of full belief, (WADA) is strictly stronger than (SADA) and (SSADA). Similarly, when it comes to comparative confidence, (WADA$_{\succeq}$) will be strictly stronger than (SADA$_{\succeq}$) and (SSADA$_{\succeq}$), both of which will also be discussed below.

[80]Note that we do not restrict this quantifier to total preorders. If there is *any* binary relation $\succeq'$ on that weakly accuracy dominates $\succeq$ on $\mathcal{A}$, then we will take this to be an *alethic defect* of $\succeq$.

[81]In fact, it can be shown that neither of the relations ($\succeq$ or $\succeq'$) discussed in this section is weakly $\mathcal{I}_{\succeq}$-dominated by *any* binary relation on $\mathcal{B}_4$. This will follow from our proof that $\mathcal{I}_{\succeq}$ is evidentially proper, below. Intuitively, this is as it should be. After all, either of these relations *could be* supported by a (rational) agent's total evidence E. For instance, $\succeq$ would be supported by E if E entailed (only) the claim that the coin is *fair*, while $\succeq'$ would be supported by E if E entailed (only) the claim that the coin is *biased toward heads* (Joyce, 2009, *pp.* 282–3). We will say more about the nature of evidentially supported comparative confidence relations in §5, below.

put this more formally (in terms of $\mathcal{I}_{\succeq}$), there should *not* exist a binary relation $\succeq'$ on $\mathcal{A}$ such that

$$(\forall w)\,[\mathcal{I}_{\succeq}(\succeq', w) < \mathcal{I}_{\succeq}(\succeq, w)].$$

And, the second is the $\succeq$-analogue of (SSADA). Let $\mathbf{M}(\succeq, w)$ denote the set of (qualitatively) *inaccurate* comparative judgments (of either of the two types discussed above) made by a comparative confidence relation $\succeq$ at a possible world w. And, consider the following *bedrock* fundamental epistemic principle.

Strong Strict Accuracy-Dominance Avoidance for Comparative Confidence Relations (SSADA$_{\succeq}$). $\succeq$ should *not be strongly strictly dominated* in accuracy. Formally (in terms of **M**), there should *not* exist a binary relation $\succeq'$ on $\mathcal{A}$ such that

$$(\forall w)\,[\mathbf{M}(\succeq', w) \subset \mathbf{M}(\succeq, w)].$$

(SADA$_{\succeq}$) entails (SSADA$_{\succeq}$), but not conversely. This is because violating (SSADA$_{\succeq}$) entails that there exists a binary relation $\succeq'$ which *not only strictly dominates* $\succeq$, but *also never makes any mistakes that* $\succeq$ *doesn't already make*. Clearly, if S's $\succeq$ violates (SSADA$_{\succeq}$), this means S is failing to live up to her epistemic aim of making accurate judgments (no matter how this is construed). Other choices of fundamental epistemic principle could be made. But, in this initial investigation, we will stick with (WADA$_{\succeq}$), (SADA$_{\succeq}$) and (SSADA$_{\succeq}$) as our three fundamental (alethic) epistemic principles.[82] In fact, mainly, we'll be making use of (WADA$_{\succeq}$). But, occasionally, we'll point out when certain claims follow from the weaker principles (SADA$_{\succeq}$) or (SSADA$_{\succeq}$). The following theorem undergirds an epistemic justification of ($\mathfrak{C}$).

THEOREM 3.1. *If $\succeq$ violates ($\mathfrak{C}$), then $\succeq$ violates (WADA$_{\succeq}$), i.e., (WADA$_{\succeq}$) $\Rightarrow$ ($\mathfrak{C}$).*

In the next section, we look beyond ($\mathfrak{C}$) to stronger coherence requirements for $\succeq$ that have appeared in the literature. As we'll see, (WADA$_{\succeq}$) can be used to provide epistemic justifications for an important family of traditional coherence requirements for $\succeq$.

4. Beyond ($\mathfrak{C}$) — A Family of Traditional Coherence Requirements for $\succeq$

The fundamental requirement ($\mathfrak{C}$) is but one among many coherence requirements that have been proposed for $\succeq$. We will not attempt to survey all of the requirements that have been discussed in the literature (Capotorti and Vantaggi, 2000). We'll focus on one particular family of requirements. Before stating the other requirements in the family, we first need to define two more types of numerical functions that will serve as *representers* of comparative confidence relations.

A *mass function* on a Boolean algebra $\mathcal{B}_n$ is real-valued function $m : \mathcal{B}_n \mapsto [0, 1]$ which maps propositions from $\mathcal{B}_n$ to the unit interval, and which satisfies the following two axioms.

(M$_1$) $m(\bot) = 0$.

[82] In §5 below, we will discuss *evidential* epistemic requirements and explain how they relate to our alethic requirements (WADA$_{\succeq}$), (SADA$_{\succeq}$) and (SSADA$_{\succeq}$).

(M_2) $\sum_{p \in \mathcal{B}_n} m(p) = 1.$

A *belief function* — sometimes called a *Dempster-Shafer function* (Dempster, 1968; Shafer, 1976; Liu and Yager, 2008) — on a Boolean algebra $\mathcal{B}_n$ is a real-valued function $\text{Bel} : \mathcal{B}_n \mapsto [0,1]$ which maps propositions from $\mathcal{B}_n$ to the unit interval, and which is generated by an underlying mass function m in the following way

$$\text{Bel}_m(p) \overset{\text{def}}{=} \sum_{\substack{q \in \mathcal{B}_n \\ q \text{ entails } p}} m(q).$$

It is easy to show that all belief functions are plausibility functions (but not conversely). In this sense, the concept of a belief function is a refinement of the concept of a plausibility function. The class of Belief functions, in turn, contains the class of *probability* functions, which can be defined in terms of a special type of mass function. Let $\mathfrak{s} \in \mathcal{B}_n$ be the *states* of a Boolean algebra $\mathcal{B}_n$ (or, if you prefer, the *state descriptions* of a propositional language $\mathcal{L}$ which generates $\mathcal{B}_n$). The states of $\mathcal{B}_n$ just correspond to the *possible worlds* that are involved in our epistemic evaluations of the agent in question. A *probability* mass function is real-valued function $\mathfrak{m} : \mathcal{B}_n \mapsto [0,1]$ which maps *states* of $\mathcal{B}_n$ to the unit interval, and which satisfies the following two axioms.

($\mathfrak{M}_1$) $\mathfrak{m}(\bot) = 0.$

($\mathfrak{M}_2$) $\sum_{\mathfrak{s} \in \mathcal{B}_n} \mathfrak{m}(\mathfrak{s}) = 1.$

A *probability function* on a Boolean algebra $\mathcal{B}_n$ is a real-valued function $\Pr : \mathcal{B}_n \mapsto [0,1]$ which maps propositions from $\mathcal{B}_n$ to the unit interval, and which is generated by an underlying probability mass function $\mathfrak{m}$ in the following way

$$\Pr_{\mathfrak{m}}(p) \overset{\text{def}}{=} \sum_{\substack{\mathfrak{s} \in \mathcal{B}_n \\ \mathfrak{s} \text{ entails } p}} \mathfrak{m}(\mathfrak{s}).$$

It is easy to show that all probability functions are belief functions (but not conversely). So, probability functions are special kinds of belief functions (and belief functions are, in turn, special kinds of plausibility measures).

There are various senses in which a real-valued function f may be said to *represent* a comparative confidence relation $\succeq$. We have already seen the strongest variety of numerical representation, which is called *full agreement*.

DEFINITION 3.1. *f **fully agrees** with a comparative confidence relation $\succeq$ just in case, for all $p, q \in \mathcal{A}$, $p \succ q$ iff $f(p) > f(q)$, and $p \sim q$ iff $f(p) = f(q)$.*

Thus, the fundamental coherence requirement ($\mathfrak{C}$) requires that there exist a plausibility measure Pl which *fully agrees* with $\succeq$. Another, weaker kind of numerical representability is called *partial agreement*.

DEFINITION 3.2. *f **partially agrees** with a comparative confidence relation $\succeq$ just in case, for all $p, q \in \mathcal{A}$, $p \succ q$ only if $f(p) > f(q)$.*

If f partially agrees with $\succeq$, then we will say that f *partially represents* $\succeq$. And, if f fully agrees with $\succeq$, then we will say that f *fully represents* $\succeq$. It is easy to see that full representability is *strictly stronger* than partial representability.

It is well known (Wong et al., 1991) that a total preorder $\succeq$ is partially represented by some belief function Bel only if $\succeq$ satisfies (A_2). The following theorem is, therefore, an immediate corollary of Theorem 3.1.

THEOREM 3.2. *($WADA_{\succeq}$) entails that $\succeq$ is (extendible to a relation on $\mathcal{B}_n$ that is — fn. 73) partially represented by some belief function Bel.*

A natural question to ask at this point is: Does ($WADA_{\succeq}$) ensure that $\succeq$ is *fully* represented by some belief function Bel? The answer is *yes*. In order to see this, it helps to recognize that full representability by a belief function has a simple axiomatic characterization (Wong et al., 1991). Specifically, a total preorder $\succeq$ is fully represented by some belief function only if $\succeq$ satisfies (A_1), (A_2), and

(A_3) For all $p, q, r \in \mathcal{A}$, if p entails q and $\langle q, r\rangle$ are mutually exclusive, then:

$$q \succ p \Longrightarrow q \vee r \succ p \vee r.$$

Theorem 3.1 establishes that ($WADA_{\succeq}$) entails both (A_1) and (A_2). Moreover, it turns out that ($WADA_{\succeq}$) is also strong enough to entail (A_3). That is, we have the following theorem regarding ($WADA_{\succeq}$).

THEOREM 3.3. *($WADA_{\succeq}$) entails (A_3). [As a result, ($WADA_{\succeq}$) entails that $\succeq$ is (extendible to a relation on $\mathcal{B}_n$ that is — fn. 73) fully represented by some belief function (Wong et al., 1991).]*

Let's take stock. So far, we have encountered the following three coherence requirements for $\succeq$, in increasing order of strength, each of which is a consequence of our fundamental epistemic principle ($WADA_{\succeq}$).

($\mathfrak{C}_0$) $\succeq$ should be partially representable by some belief function Bel. This is equivalent to requiring that $\succeq$ (a total preorder) satisfies (A_2).

($\mathfrak{C}$) $\succeq$ should be fully representable by some plausibility measure Pl. This is equivalent to requiring that $\succeq$ (a total preorder) satisfies (A_1) and (A_2).

($\mathfrak{C}_1$) $\succeq$ should be fully representable by some belief function Bel. This is equivalent to requiring that $\succeq$ (a total preorder) satisfies (A_1), (A_2), and (A_3).

Moving beyond ($\mathfrak{C}_1$) takes us into the realm of *comparative probability*. A total preorder $\succeq$ is said to be a *comparative probability* relation only if $\succeq$ satisfies (A_1) and the following two additional axioms.

(A_4) For all $p \in \mathcal{A}$, $p \succeq \bot$.

(A_5) For all $p, q, r \in \mathcal{A}$, if $\langle p, q\rangle$ are mutually exclusive and $\langle p, r\rangle$ are mutually exclusive, then both of the following biconditionals are true:

$$q \succ r \iff p \vee q \succ p \vee r$$
$$\&$$
$$q \sim r \iff p \vee q \sim p \vee r$$

It is easy to show that {(A_1), (A_2)} jointly entail (A_4). So, $\succeq$ (a total preorder) is a comparative probability relation just in case $\succeq$ satisfies the three axioms (A_1), (A_2) and (A_5). Now, consider the following coherence requirement.

($\mathfrak{C}_2$) $\succeq$ should be be a comparative probability relation. This is equivalent to requiring that $\succeq$ (a total preorder) satisfies (A_1), (A_2) and (A_5).

It is well known (and not too difficult to prove) that (A_5) is *strictly stronger* than (A_3), in the presence of (A_1) and (A_2). Therefore, ($\mathfrak{C}_2$) is *strictly stronger* than ($\mathfrak{C}_1$). Moreover, in light of the following negative result, ($\mathfrak{C}_1$) is also where ($\text{WADA}_{\succeq}$)'s coherence implications peter out.

THEOREM 3.4. *($\text{WADA}_{\succeq}$) does not entail (A_5). In light of Theorems 3.1 and 3.3, this is equivalent to the claim that ($\text{WADA}_{\succeq}$) does not entail ($\mathfrak{C}_2$).*

Theorem 3.4 reveals that — just as in the case of full belief above — weak accuracy dominance avoidance yields a coherence requirement that is *weaker than (full) probabilistic representability* of the judgment set. Other analogies (and disanalogies) between the full belief case and the comparative confidence case will be discussed below (and in the Negative Phase). Meanwhile, we will examine a few more traditional coherence requirements for $\succeq$.

de Finetti (1937, 1951) famously conjectured that all comparative probability relations are fully representable by some probability function. As it turns out, this conjecture is false. In fact, Kraft et al. (1959) showed that some comparative probability relations are *not even partially* representable by any probability function. As a result, the following coherence requirement for $\succeq$ is *strictly stronger* than ($\mathfrak{C}_2$).

($\mathfrak{C}_3$) $\succeq$ should be be partially representable by some probability function.

And, of course, ($\mathfrak{C}_3$) is *strictly weaker* than the following coherence requirement, which is the strongest of all the traditional coherence requirements.

($\mathfrak{C}_4$) $\succeq$ should be be fully representable by some probability function.

Moreover, it is easy to show (a) that ($\mathfrak{C}_3$) is *independent* of both ($\mathfrak{C}_2$) and ($\mathfrak{C}_1$), and (b) that ($\mathfrak{C}_3$) is *strictly stronger* than ($\mathfrak{C}$). The following axiomatic constraint is a slight weakening of (A_5).

($A_5^\star$) For all $p, q, r \in \mathcal{A}$, if $\langle p, q\rangle$ are mutually exclusive and $\langle p, r\rangle$ are mutually exclusive, then:

$$q \succ r \implies p \vee r \not\succ p \vee q$$

And, the following coherence requirement is a (corresponding) weakening of ($\mathfrak{C}_2$).

($\mathfrak{C}_2^\star$) $\succeq$ should (be a total preorder and) satisfy (A_1), (A_2) and ($A_5^\star$).

Finally, the following negative result shows that ($\mathrm{WADA}_{\succeq}$) is not strong enough to entail even *partial* probabilistic representability of $\succeq$.

Theorem 3.5. *($\mathrm{WADA}_{\succeq}$) does not entail ($\mathfrak{C}_2^\star$). [And, as a result, ($\mathrm{WADA}_{\succeq}$) does not entail ($\mathfrak{C}_3$) either.]*

The arrows in Figure 3.3 depict the logical relations between the traditional coherence requirements we've been discussing here. The superscripts on the coherence requirements in Figure 3.3 have the following meanings. If a coherence requirement is known to follow from ($\mathrm{WADA}_{\succeq}$), then it gets a "$\checkmark$". And, if a coherence requirement is known to follow from ($\mathrm{SSADA}_{\succeq}$), then it gets a "$\checkmark\checkmark$." If a coherence requirement is known *not* to follow from ($\mathrm{WADA}_{\succeq}$), then it gets an "**✗**".

All three "$\checkmark$"s (and the "$\checkmark\checkmark$") in Figure 3.3 are established by our proof of Theorem 3.1. The "**✗**" on ($\mathfrak{C}_2$) is established by our proof of Theorem 3.4 and the "**✗**"s on ($\mathfrak{C}_2^\star$) and ($\mathfrak{C}_3$) are established by our proof of Theorem 3.5.

$$
\begin{tikzcd}
\text{full rep. by Pr} & (\mathfrak{C}_4)^{\text{✗}} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & (\mathfrak{C}_3)^{\text{✗}} \arrow[d, Rightarrow] & \text{partial rep. by Pr} \\
\text{qualitative prob.} & (\mathfrak{C}_2)^{\text{✗}} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & (\mathfrak{C}_2^\star)^{\text{✗}} \arrow[ldd, Rightarrow] & (\mathrm{A}_1) + (\mathrm{A}_2) + (\mathrm{A}_5^\star) \\
\text{full rep. by Bel} & (\mathfrak{C}_1)^{\checkmark} \arrow[d, Rightarrow] & & \\
\text{full rep. by Pl} & (\mathfrak{C})^{\checkmark} \arrow[d, Rightarrow] \arrow[rd, Rightarrow] & & \\
\text{partial rep. by Bel} & (\mathfrak{C}_0)^{\checkmark} & (\mathrm{A}_1)^{\checkmark\checkmark} & \top \succ \bot
\end{tikzcd}
$$

Figure 3.3. Logical relations between traditional $\succeq$-coherence requirements

5. Evidential Requirements for Comparative Confidence

So far, our approach has been based on *alethic* considerations. As in the case of full belief, alethic requirements do not seem to be the only kinds of epistemic requirements for comparative confidence. Indeed, we have already seen an example of a comparative confidence relation that seems to be epistemically reasonable/rational, but which *must* be *in*accurate. That example is depicted in Figure 3.1. It involves an agent who is indifferent between P and $\neg P$, where P is the claim that a coin which has been tossed has landed heads. Let's suppose, further, that the agent's total evidence entails that the coin in question is *fair* (and nothing else). Because P and $\neg P$ cannot have the same truth-value, our definition of inaccuracy entails that $P \sim \neg P$ *must* be inaccurate (*viz.*, alethically defective). Nonetheless, intuitively, the judgment $P \sim \neg P$ is *supported by the agent's total evidence.* It is for this reason that we deem the judgment in question to be epistemically rational/reasonable. At this point, we need to say something about the *general* evidential requirements for comparative confidence.

There are various proposals one could make regarding the (general) evidential requirements for comparative confidence. For the purposes of this investigation (and in keeping with the applications of our framework to full belief and numerical credence), we will adopt a *probabilistic* approach to evidential support.

In the case of full belief, we were working with an understanding of *probabilistic evidentialism* (recall its slogan "probabilities reflect evidence") that may be summarized *via* the following, general *necessary condition* for a (qualitative) judgment's being "supported by the total evidence".

($\mathcal{E}$) A qualitative judgment j (of type $\mathfrak{J}$) is supported by the total evidence *only if* there exists some probability function that probabilifies (*i.e.*, assigns probability greater than $1/2$ to) the vindicator (or "accuracy-maker") of j.

The vindicator of a belief that p [$B(p)$] is p and the vindicator of a disbelief that p [$D(p)$] is $\neg p$. The general principle ($\mathcal{E}$) thus led us to the following necessary condition for a belief set **B** (on an agenda $\mathcal{A}$) to be supported by the evidence.

Necessary Condition for Satisfying the Evidential Requirement for Full Belief. A full belief set **B** satisfies the evidential requirement, *i.e.*, all judgments in **B** are supported by the total evidence, ***only** if*

($\mathcal{R}$) There exists some probability function that probabilifies, *i.e.*, assigns probability greater than $1/2$ to:

(i) p, for each belief $B(p) \in$ **B**,

(ii) $\neg p$, for each disbelief $D(p) \in$ **B**.

In the case of comparative confidence, we have a set of $\binom{m}{2}$ comparative confidence judgments over some m-proposition agenda $\mathcal{A}$. Some of these judgments may be strict comparisons $p \succ q$ and some may be indifferences $p \sim q$. Initially (*fn.* 74), it may seem natural to say that "the vindicator of" a strict comparison $p \succ q$ is $p \mathbin{\&} \neg q$ and that "the vindicator of" an indifference $p \sim q$ is $p \equiv q$ (*fn.* 74). On this (*prima facie* natural) way of thinking (which, as we'll see shortly, is *flawed*), and in keeping with ($\mathcal{E}$), the corresponding (putative) necessary condition for satisfying the evidential requirement for comparative confidence would seem to be

Putative Necessary Condition for Satisfying the Evidential Requirement for Comparative Confidence. A comparative confidence judgment set **C** satisfies the evidential requirement, *i.e.*, all judgments in the set are supported by the total evidence, ***only** if*

There exists some probability function that probabilifies, *i.e.*, assigns probability greater than $1/2$ to:

(i) $p \mathbin{\&} \neg q$, for each strict judgment $p \succ q \in$ **C**,

(ii) $p \equiv q$, for each indifference judgment $p \sim q \in$ **C**.

This *putative* evidential requirement is *not* a *bona fide* requirement. In fact, neither clause (i) nor clause (ii) expresses a genuine evidential requirement for comparative confidence relations. We have already seen why clause (ii) won't do. If we require that $p \equiv q$ be probabilified relative to some probability function, then this will rule out $P \sim \neg P$ as evidentially unsupportable, since $\Pr(P \equiv \neg P) = 0$, on *every* probability function. However, our fair coin toss example illustrates that the evidential requirements entail no such thing. The fair coin toss example also suffices to explain why (i) fails to express a genuine evidential requirement. In this case, we have $P \succ \bot$ and $\neg P \succ \bot$. And, intuitively, both of these judgments would be supported by the agent's total evidence. However, there can be no probability function that probabilifies the "vindicators" of both of these judgments simultaneously. That is, there can be no probability function that simultaneously assigns probability greater than $1/2$ to both P and $\neg P$. Therefore, clause (i) does not express a genuine (probabilistic) evidential requirement for comparative confidence either.

These considerations reveal an important disanalogy between full belief and comparative confidence. In the case of full belief, the "master (qualitative) probabilistic evidentialist" principle $(\mathcal{E})$ leads to a *bona fide* (probabilistic) evidential requirement [*viz.*, $(\mathcal{R})$]. But, in the case of comparative confidence, an apparently analogous application of $(\mathcal{E})$ implies conditions that do *not* correspond to *bona fide* (probabilistic) evidential requirements. The problem here is that the application of $(\mathcal{E})$ to the case of individual comparative confidence judgments is *not* analogous (or appropriate). Strictly speaking, given our approach to Step 1 (*fn.* 74), there is no such thing as "the vindicator of" an individual comparative confidence judgment. There are (sometimes non-unique) vindicated *total comparative confidence ordering*(s). These are the ordering(s) which rank all the truths strictly above all the falsehoods (at the possible world in question). When it comes to *individual* comparative confidence judgments, all we can say is that there are two ways in which they can be *in*accurate — the two ways identified in Facts 1 and 2 above. We will return to these issues in the Negative Phase.

In any case, here is a natural alternative to the (putative, probabilistic) evidential requirement implied by $(\mathcal{E})$ in the case of comparative confidence.

> **Necessary Condition for Satisfying (Probabilistic) Evidential Requirements for Comparative Confidence**. A comparative confidence relation $\succeq$ (on an agenda $\mathcal{A}$) satisfies the (probabilistic) evidential requirement, *i.e.*, all of $\succeq$'s judgments are supported by the total evidence, ***only if***:

$(\mathcal{R}_{\succeq})$ There exists some regular[83] probability function $\Pr(\cdot)$ such that:

(i) $\Pr(p) > \Pr(q)$, for each $p, q \in \mathcal{A}$ such that $p \succ q$,

(ii) $\Pr(p) = \Pr(q)$, for each $p, q \in \mathcal{A}$ such that $p \sim q$.

Of course, $(\mathcal{R}_{\succeq})$ entails the requirement that $\succeq$ be fully representable by some probability function, which was the strongest of the historical coherence requirements

[83] A probability function is *regular* just in case it only assigns extreme probabilities to logically non-contingent claims. Equivalently, a probability function is regular only if it assigns non-zero probability to all logically possible worlds (Shimony, 1955). Recall (*fn.* 70), we are assuming that our agents have Regular comparative confidence orderings. Moreover, the assumption of numerical regularity will be important to ensure the logical connection between $(\mathcal{R}_{\succeq})$ and $(\text{WADA}_{\succeq})$. We explain why numerical regularity is needed for these purposes in the proof of Theorem 3.6 in Appendix C, and we discuss some philosophical ramifications of numerical regularity in the Negative Phase.

that we discussed above [*viz.*, ($\mathcal{R}_{\succeq}$) $\Rightarrow$ ($\mathfrak{C}_4$)]. The following theorem provides the key theoretical connection between alethic and evidential requirements for $\succeq$.

THEOREM 3.6. *If a comparative confidence relation $\succeq$ satisfies ($\mathcal{R}_{\succeq}$), then $\succeq$ satisfies (WADA$_{\succeq}$). That is, ($\mathcal{R}_{\succeq}$) entails (WADA$_{\succeq}$).*

This theorem is analogous to to our central Theorem 1.2 from Part I. And, the proof of Theorem 3.6 is analogous to the proof of Theorem 1.2, insofar as it reveals (a) that any comparative confidence relation that is fully representable by some (regular) probability function Pr will also *minimize expected $\mathcal{I}_{\succeq}$-inaccuracy*, relative to its representing (evidential) probability function Pr; and, as a result, that (b) the measure $\mathcal{I}_{\succeq}$ is *evidentially proper*.[84] One final corollary of this central theorem for comparative confidence is that the entailment ($\mathcal{R}_{\succeq}$) $\Rightarrow$ (WADA$_{\succeq}$) will hold even if we assign different weights to different pairs of propositions in the agenda $\mathcal{A}$ (in our calculation of the total $\mathcal{I}_{\succeq}$-score of a comparative confidence judgment set). So, although our measure $\mathcal{I}_{\succeq}$ is additive, its use is compatible with assigning different "importance weights" to different pairs of propositions.

In the next section, we will step back and look at the big picture. We'll see how the alethic requirements and the evidential requirements for comparative confidence fit together, and how they dovetail with existing coherence requirements.

6. The Big Picture

From an alethic point of view, there are various norms which may govern comparative confidence judgments. In addition to (WADA$_{\succeq}$), there are also stronger alethic norms that one might consider. The strongest of these alethic norms would require that all comparative confidence judgments are (actually) *not in*accurate (in either of the ways identified in Facts 1 or 2).

> **Actual Vindication for Comparative Confidence Judgments** (AV$_{\succeq}$). No member of any set of comparative confidence judgments $\mathbf{C}$ should be *actually* inaccurate (*i.e.*, if $p \succ q \in \mathbf{C}$, then $p \vee \neg q$ should be *actually* true, and if $p \sim q \in \mathbf{C}$, then $p \equiv q$ should be *actually* true).

(AV$_{\succeq}$) is the analogue of the truth norm (TB) for full belief (which just is the requirement of actual vindication for full belief). A weaker alethic requirement would require only *possible* vindication for comparative confidence judgments.

> **Possible Vindication for Comparative Confidence Judgments** (PV$_{\succeq}$). For each set of comparative confidence judgments $\mathbf{C}$, there should be some *possible* world at which no member of $\mathbf{C}$ is inaccurate (*i.e.*, there should be some world w such that (a) for each p, q such that $p \succ q \in \mathbf{C}$, $p \vee \neg q$ is true at w, and (b) for each p, q such that $p \sim q \in \mathbf{C}$, $p \equiv q$ is true at w).

(PV$_{\succeq}$) is the analogue of the deductive consistency requirement for full belief (which just is the requirement of possible vindication for full belief).

On the evidential side, we have the following (master) evidential requirement for comparative confidence judgments.

[84] We will also prove (in Appendix C, along with our proof of Theorem 3.6) that the particular choice of inaccuracy scores (*i.e.*, the 2:1 ratio of scores in $\acute{\imath}_{\succeq}$) for the two types of mistakes identified in Facts 1 and 2 above is *forced* by the requirement that the inaccuracy measure be evidentially proper.

> **Evidential Requirement for Comparative Confidence Judgments** ($\mathrm{E}_{\succeq}$). For each set of comparative confidence judgments **C**, all of **C**'s members should be supported by the total evidence.

From a *probabilistic evidentialist* perspective, ($\mathrm{E}_{\succeq}$) requires (at least[85]) that $\succeq$ satisfy condition ($\mathcal{R}_{\succeq}$) above. At this point, we have all the ingredients we need to reveal the big picture of norms and requirements discussed in this part of the book.

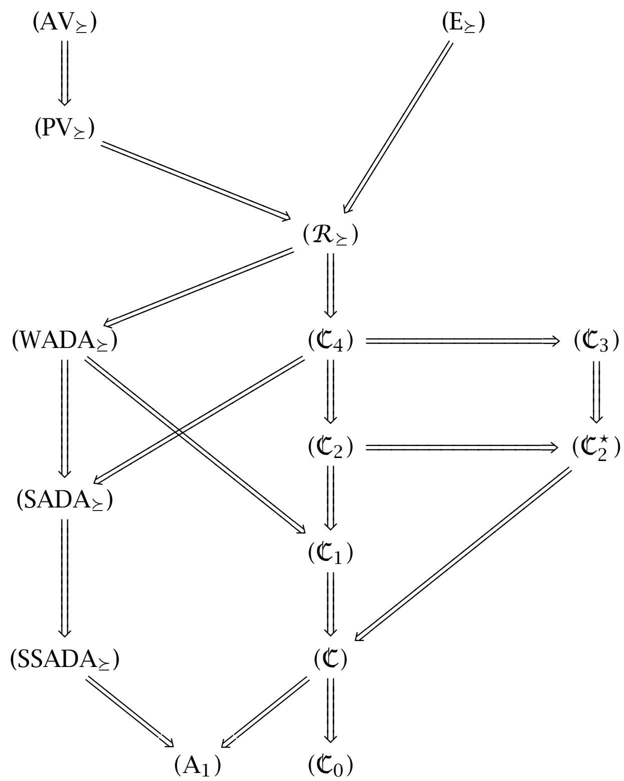

FIGURE 3.4. Logical relations between all norms and requirements for comparative confidence

Figure 3.4 reveals that (as in the case of full belief) our requirement ($\mathcal{R}_{\succeq}$) plays a central role in our story. For ($\mathcal{R}_{\succeq}$) is entailed by both of the strong alethic requirements ($\mathrm{AV}_{\succeq}$) and ($\mathrm{PV}_{\succeq}$), and the (master) evidential requirement ($\mathrm{E}_{\succeq}$).[86] Of course, ($\mathcal{R}_{\succeq}$) entails all of the traditional coherence requirements ($\mathfrak{C}_i$) discussed above. Moreover, as the proof of Theorem 3.6 reveals, ($\mathcal{R}_{\succeq}$) also entails ($\mathrm{WADA}_{\succeq}$).

---

[85] Intuitively, just as in the case of full belief, the evidential requirement ($\mathrm{E}_{\succeq}$) will be *stronger* than ($\mathcal{R}_{\succeq}$), since it will require representability by a *specific* (regular) *evidential* probability function, which gauges the evidential support relations in the context in which the $\succeq$-judgments are formed.

[86] Because all of our comparative confidence relations are *regular* total preorders, we should always be able to construct a regular probability function $\Pr(\cdot)$ which represents a comparative confidence $\succeq$ relation that is "possibly vindicated" in the sense of ($\mathrm{PV}_{\succeq}$) in possible world w. I haven't been able to specify a general algorithm for such a construction, but I'm confident that there must be such a construction. Assistance welcome!

Finally, our arguments above reveal that ($\text{WADA}_{\succeq}$) implies that an agent's comparative confidence ordering should be (extendible to a relation on $\mathcal{B}_n$ that is — *fn.* 73) fully representable by some belief function [($\mathfrak{C}_1$)]. This provides an epistemic justification for the "coherence" of comparative confidence judgments in the sense of Dempster-Shafer (Dempster, 1968; Shafer, 1976; Liu and Yager, 2008). Adherents of representability by a belief function as a requirement of rationality should welcome this result — especially, since we know of no existing epistemic justification for ($\mathfrak{C}_1$) as a formal coherence requirement for comparative confidence.

CHAPTER 4

Comparative Confidence: Negative Phase

What?

Richard Milhous Nixon

1. Why Not the Stronger Notion of "Accuracy" for Individual $\succeq$ Judgments?

Recall the assumption (which we discussed above — *fn.* 74) that there is a *unique* vindicated ordering $\mathring{\succeq}_w$ in each world w, which is defined as follows

$$p \mathrel{\mathring{\succeq}_w} q \overset{\text{def}}{=} \begin{cases} p \mathrel{\mathring{\succ}_w} q & \text{if } p \text{ is true in } w \text{ and } q \text{ is false in } w, \\ p \mathrel{\mathring{\sim}_w} q & \text{if } p \text{ and } q \text{ have the same truth-value in } w. \end{cases}$$

This assumption leads to a weaker notion of "*in*accuracy of a comparative confidence ordering $\succeq$" — *non-identity* of $\succeq$ and $\mathring{\succeq}_w$ — than our (official) definition, which is compatible with the existence of *many* "vindicated orderings" in a given world w — so long as they each rank all the truths strictly above all the falsehoods at w. This stronger notion of "accuracy" suggests other possibilities for non-zero inaccuracy scores that should be assigned by our point-wise measure $\mathfrak{i}_{\succeq}$. Figure 4.1 depicts all the possible point-wise scoring schemes for individual comparative confidence judgments. Here, we stipulate (as a matter of convention) that

w_i	p	q	$p \succ q$	$p \sim q$
w_1	T	T	b	0
w_2	T	F	0	1
w_3	F	T	a	1
w_4	F	F	b	0

TABLE 4.1. All possible point-wise scoring schemes for individual comparative confidence judgments.

the inaccuracy score for an indifference judgment regarding propositions having different truth values is 1. The zeros in the table are uncontroversial, assuming that our measure is $\mathfrak{i}$ is going to be *truth-directed* (*fn.* 30). The only two parameters that remain are a and b. Our (stronger) notion of inaccuracy for individual comparative confidence judgments — as embodied in our definition of $\mathfrak{i}_{\succeq}$ — implies that $b = 0$. That's because these judgments do not (in and of themselves) entail that the relation $\succeq$ is inaccurate in our official sense. However, these judgments *would* imply that $\succeq$ is inaccurate in the (weaker) sense above (*viz.*, *being non-identical with* $\mathring{\succeq}_w$). Unfortunately, *no point-wise measure of inaccuracy that assigns anything other than* $a := 2$ *and* $b := 0$ *leads to an evidentially proper measure of total inaccuracy.*[87] For this reason, we *must* assign $b := 0$ if we want an evidentially proper measure of inaccuracy to emerge from our approach. This is the main reason we have opted for our weaker notion of "accuracy" (and our stronger notion of "inaccuracy").[88] In the next section, we examine a formal analogy between full belief and comparative confidence that has been explored in the literature.

2. A Formal Analogy Between Full Belief and Comparative Confidence: Some Cautionary Remarks

It is sometimes argued (Duddy and Piggins, 2012) that it is possible to construct an adequate measure $\delta(\succeq, w)$ of the inaccuracy of a comparative confidence relation $\succeq$ by (a) "translating" its comparative confidence judgments $p \succeq q$ into *full beliefs* that some (extensional) "is at least as plausible as" relation obtains between p and q, and (b) applying one's favored measure of inaccuracy for belief sets to these "translated" comparative confidence judgment sets. Specifically, we can use our (evidentially proper) measure of the inaccuracy of a belief set $\mathcal{I}(\mathbf{B}, w)$ to construct a measure $\delta(\succeq, w)$ of the inaccuracy of comparative confidence relations.

Here's how the construction works. For each pair of propositions $p, q \in \mathcal{A}$, our agents make exactly one comparative judgment out of the following three

$$\{p \succ q, q \succ p, p \sim q\}.$$

If we define the relation:

$$p \succcurlyeq q \overset{\text{def}}{=} (p \succ q) \text{ or } (p \sim q),$$

then these three judgments correspond to the following three $\succcurlyeq$-judgment *pairs*

$$\{\langle p \succcurlyeq q, q \not\succcurlyeq p\rangle, \langle p \not\succcurlyeq q, q \succcurlyeq p\rangle, \langle p \succcurlyeq q, q \succcurlyeq p\rangle\}.$$

Next, we can "translate" these $\succcurlyeq$-pairs into pairs of *beliefs* and *disbeliefs*, as follows

$$\{\langle B(Rpq), D(Rqp)\rangle, \langle D(Rpq), B(Rqp)\rangle, \langle B(Rpq), B(Rqp)\rangle\},$$

where Rpq is interpreted as some (extensional) "at least as plausible as" relation that may obtain between propositions p and q. To complete the analogy, we just need an extensional (*viz.*, truth-functional) interpretation of the relation R. Given our definition of inaccuracy for comparative confidence judgments $p \succeq q$, the appropriate (extensional) interpretation of $\ulcorner Rpq \urcorner$ is $\ulcorner p \vee \neg q \urcorner$.

We can now construct a measure of inaccuracy δ for (sets of) comparative confidence judgments in terms of our naïve mistake-counting measure of inaccuracy

[87] This is proven in Appendix C (after the proof of Theorem 3.6).

[88] We conjecture that it is possible to think of our measure $\mathcal{I}_{\succeq}$ as a measure of "distance from vindication", even though there may be *many* "vindicated relations" $\{\mathring{\succeq}_w\}$ in a given possible world w. The idea would be to define $\succeq$'s "distance from vindication at w" as the *edit distance* (Deza & Deza 2009) between $\succeq$ and the member of $\{\mathring{\succeq}_w\}$ that is closest to $\succeq$ (in edit distance). We thank Ben Levinstein for suggesting this conjecture. This is another interesting open theoretical question.

$\mathcal{I}$ for belief sets, *via* the above translation. First, for each comparative confidence judgment in a set $\mathbf{C}$, we translate it into its corresponding belief/disbelief pair. Then, we combine all these pairs into one big belief set $\mathbf{B}_{\mathbf{C}}$. In this way, every comparative confidence judgment set $\mathbf{C}$ will have a (unique) corresponding belief set $\mathbf{B}_{\mathbf{C}}$. Finally, applying our naïve mistake-counting measure of inaccuracy for belief sets to $\mathbf{B}_{\mathbf{C}}$ yields $\mathcal{I}(\mathbf{B}_{\mathbf{C}}, w)$. It is easy to show (Duddy and Piggins, 2012) that this procedure yields an inaccuracy measure $\delta(\mathbf{C}, w) \stackrel{\text{def}}{=} \mathcal{I}(\mathbf{B}_{\mathbf{C}}, w)$ for comparative confidence judgment sets, which is equivalent to the measure that simply *counts the number of differences between the adjacency matrix of* $\mathbf{C}$ *and the adjacency matrix of "the unique vindicated* $\succeq$*-relation at* w*"* ($\mathring{\succeq}_w$) that we discussed in the previous section (*fn.* 74). In other words, $\delta(\mathbf{C}, w)$ is the *Kemeny distance* (Deza & Deza 2009) between the two total preorders $\succeq$ (*viz.*, $\mathbf{C}$) and $\mathring{\succeq}_w$.

Because the Kemeny measure — like our simple, mistake-counting measure Δ above — simply *counts* "the number of mistakes" in a comparative judgment set at a world, it *scores all mistakes equally*. As a result, we have the following negative result regarding the Kemeny measure of inaccuracy of a $\succeq$-relation.

THEOREM 4.1. *The entailment* ($\mathcal{R}_{\succeq}$) $\Rightarrow$ (WADA$_{\succeq}$) *does not hold, if we use the Kemeny measure of inaccuracy* (δ) *rather than our (evidentially proper) measure* $\mathcal{I}_{\succeq}$. *And, as a result, the Kemeny measure of inaccuracy is not evidentially proper (even for regular evidential probability functions).*

So, like Δ, the Kemeny measure runs afoul of our evidential requirements for comparative confidence. This is interesting, as it shows that the standard way of defining the inaccuracy of a comparative confidence relation, *via* "translation" into a corresponding full belief set *fails to preserve the property of evidential propriety* that is enjoyed by our measure of inaccuracy $\mathcal{I}$ for full belief sets.

The proof of Theorem 4.1 (Appendix C) requires examples involving algebras containing at least three states (or worlds). But, there is a simpler way to see that this formal analogy between comparative confidence and full belief can (if taken too literally) yield incorrect conclusions about evidential requirements. Consider the simple agenda $\mathcal{A} = \{P, \neg P\}$. And, consider the comparative confidence relation on $\mathcal{A}$ such that $P \sim \neg P$. This corresponds to the singleton comparative confidence judgment set $\mathbf{C} = \{P \sim \neg P\}$. Now, if we "translate" $\mathbf{C}$ into its corresponding belief set, we get: $\mathbf{B}_{\mathbf{C}} = \{B(P \vee P), B(\neg P \vee \neg P)\}$, which, by logical omniscience, is just the contradictory pair of beliefs $\mathbf{B}_{\mathbf{C}} = \{B(P), B(\neg P)\}$. Neither $\mathbf{C}$ nor $\mathbf{B}_{\mathbf{C}}$ can be vindicated (in any possible world). In this (formal, alethic) sense, the analogy between $\mathbf{B}_{\mathbf{C}}$ and $\mathbf{C}$ is tight. But, as we discussed above, because it violates ($\mathcal{R}$), the pair of beliefs in $\mathbf{B}_{\mathbf{C}}$ *cannot both be supported by any body of evidence*. However, the original comparative confidence judgment $P \sim \neg P$ comprising $\mathbf{C}$ *can* be supported by some bodies of evidence (*e.g.*, our fair coin case above). The moral of this story is that we must not take formal (alethic) analogies between comparative confidence and full belief too seriously, since this will inevitably lead to inappropriate (analogical) epistemic conclusions regarding evidential requirements.[89]

[89]Before settling on the present way of modeling comparative confidence judgments and their coherence requirements, we attempted to model comparative confidence judgments *as* beliefs about comparative plausibility relations. That approach ultimately failed to yield plausible results, and was plagued by puzzles and paradoxes (not unrelated to the cautionary remarks of this section). We are

3. Regularity and Infinite Underlying Possibility Spaces

In the case of full belief, the (naïve, mistake-counting) measure of doxastic inaccuracy we discussed only makes sense as applied to finite agendas. However, our probabilistic coherence requirements ($\mathcal{R}_r$) can be non-trivially applied to infinite agendas. Of course, there are many controversies involving probability models over infinite structures (especially, uncountable ones — see below). But, in the full belief case, we can largely set those aside, and apply our favorite account of numerical (evidential) probability (for infinite epistemic possibility spaces) to give substance to our requirements ($\mathcal{R}_r$), as applied to infinite agendas.

In the case of comparative confidence, however, there are some additional wrinkles that are worth discussing here. Generally, it seems rationally permissible to rank a contingent claim P (which is not known to be false) *strictly above* a contradiction in one's comparative confidence ordering (indeed, we are assuming comparative Regularity, which *entails* this). But, some contingent claims involving an uncountable set of (underlying, epistemic) possibilities also seem to have a numerical (evidential) probability of (identically) *zero*. This combination is troublesome, given our present approach. To make things concrete, let P be the claim that a fair coin will repeatedly land heads *infinitely often* (if tossed an infinite number of times).[90] We are inclined to agree with recent authors (Williamson, 2007; Hájek, 2012; Easwaran, 2014) who argue for the following two claims.[91]

(1) It is rationally permissible to have a comparative confidence ordering $\succeq$ such that $P \succ \bot$ (indeed, we're assuming $\succeq$-*regularity*, which *entails* this).

(2) All (evidential) probability functions are such that $\Pr(P) = \Pr(\bot) = 0$.

Unfortunately, (1) and (2) jointly *contradict* the claim that ($\mathcal{R}_{\succeq}$) is a requirement of epistemic rationality. That is, the idea that "(numerical) probabilities reflect evidence" seems to run into trouble when applied to examples such as these.

Such examples are very subtle, and a proper discussion of them is beyond our present scope. However, I will make a few brief remarks about such cases. First, the agenda in question is actually *finite* — it involves just two propositions: $\{P, \bot\}$. So, we can apply our naïve framework directly to such examples, and what we will find is that, in general, $P \succ \bot$ will be *non-dominated* in accuracy. So, in support of (1), at least we can say that $P \succ \bot$ doesn't violate (WADA$_{\succeq}$). Hence, the trouble here seems to lie squarely with claim (2). It is only when we accept (2) that these examples cause problems for our overall approach, which places ($\mathcal{R}_{\succeq}$) at center stage. In other words, these cases put pressure on the claim that ($\mathcal{R}_{\succeq}$) is a (universal) requirement of epistemic rationality. In light of these examples, we are

---

indebted to David Christensen, Ben Levinstein and Daniel Berntson for pressing on these disanalogies and nudging us toward our new approach. See (Levinstein, 2013, Chapter 1) for further discussion.

[90] If you don't like this example, we could let P be the claim that an infinitely thin dart, which is thrown at the real unit interval, lands *exactly* on (say) the point $1/2$ (Hájek, 2012).

[91] Strictly speaking, most authors argue for a *credal* rendition of claim (2). That is, most authors focus on whether claim (2) is true for *epistemically rational credences*. Some authors have voiced similar reservations regarding *chances* (Barrett, 2010). I suppose it is *possible* that *evidential* probabilities could violate (2), even if rational credences (and chances) do not. But, this seems implausible.

inclined to restrict the domain of application of the present models [*viz.*, ($\mathcal{R}_{\succeq}$)] to cases involving finite underlying epistemic possibility spaces.[92]

Another important upshot of these sorts of examples is that they call into question the common assumption that comparative confidence *reduces to* numerical credence (Easwaran, 2014). Indeed, these kinds of examples seem to suggest that comparative confidence cannot even *supervene* on numerical credence (at least, not in the usual way, *via* full probabilistic representability of the comparative confidence ordering by the agent's numerical credence function). Such a failure of supervenience of comparative confidence on numerical credence would support a *pluralistic* stance regarding these two kinds of attitudes. While we take no (official) stand on this issue here, we are (as we mentioned above) inclined toward a pluralistic stance (quite generally) anyhow. So this result would not be unwelcome.

4. The Ordering Assumptions (Opinionation and Transitivity) Revisited

For the purposes of the Positive Phase, we presented a simplified version of the framework in which $\succeq$ is assumed to be a total preorder. However, three of our ordering assumptions ($\succeq$-opinionation, $\succ$-transitivity and $\sim$-transitivity) have been the source of considerable controversy in the literature (Forrest, 1989; Fishburn, 1986; Lehrer and Wagner, 1985).

Our assumption of $\succeq$-opinionation as a constraint on $\succeq$ is analogous to our assumption of opinionation in the case of full belief. For simplicity, this monograph is (almost entirely) concerned with opinionated judgment sets. As a result, I will not discuss $\succeq$-opinionation further, since that would get us into a more general discussion of suspension-like attitudes, which is beyond our present scope.

The more interesting ordering assumptions (for present purposes) are our two transitivity assumptions. Interestingly, *neither* of our transitivity assumptions follows from (WADA$_{\succeq}$). Indeed, we can say something even stronger than this. Neither of the following two forms of transitivity is implied by (WADA$_{\succeq}$).

Weak $\succ$-Transitivity. For all $p, q, r \in \mathcal{A}$, *if* $p \succ q$ and $q \succ r$, *then* $r \nsucc p$.

Transitivity of $\sim$. For all $p, q, r \in \mathcal{A}$, *if* $p \sim q$ and $q \sim r$, *then* $p \sim r$.

Weak $\succ$-Transitivity is sometimes called *weak consistency*, and it seems to be the least controversial of the varieties of $\succeq$-transitivity. Transitivity of $\sim$ seems to be the most controversial of the varieties of $\succeq$-transitivity (Lehrer and Wagner, 1985). Here is a potential counterexample to the transitivity of epistemic indifference.

> S observes a bank robbery. S gets a good look at the robber (r_0), who has a full head of hair. The police create n perfect duplicates of the robber (r_i). They remove i hairs from the head of r_i, and make a *line-up*: $r_0, r_1, r_2, r_3, \ldots, r_N$. They show S *only neighboring pairs*: $\langle r_0, r_1\rangle$, $\langle r_1, r_2\rangle$, $\langle r_2, r_3\rangle, \ldots, \langle r_{N-1}, r_N\rangle$. Now, let $p_i \overset{\text{def}}{=} r_i$ is the robber (*i.e.*, $r_i = r_0$). On the basis of the visual evidence S obtains from observing the pairs, S comes form the pairwise indifference judgments

[92] Alternatively, we could plump for an approach to epistemic decision theory that does not rely so heavily on numerical representation theorems, and places more emphasis on dominance-style reasoning — especially in cases involving uncountable underlying epistemic possibility spaces (Easwaran, 2013a). Note: we are not optimistic about the prospects of using non-standard analysis to handle such problems. See (Pruss, 2012, 2013; Easwaran, 2014) for some pessimistic arguments.

> $p_i \sim p_{i+1}$, for each i. But, at the end of the day, the police show S *one last pair*: $\langle r_0, r_N \rangle$, where r_N has *zero* hairs on his head. Of course, S judges $p_0 \succ p_N$.

In light of such examples, the fact that transitivity of indifference does not have an accuracy-dominance rationale should (perhaps) not be terribly surprising. But, it is interesting that even the least controversial variety of transitivity (weak $\succ$-transitivity) fails to have an accuracy dominance rationale.[93]

The present discussion raises a crucial theoretical question: *Precisely which* properties of $\succeq$ follow from (WADA$_{\succeq}$)? Ideally, it would be nice to have an (sound and complete) *axiomatization* of the formal properties of $\succeq$ that are consequences of (WADA$_{\succeq}$). This remains an important unsolved theoretical problem.

5. The "Circularity Worry" About ($\mathcal{R}_{\succeq}$) and Our Philosophical Methodology

One might worry that adopting ($\mathcal{R}_{\succeq}$) as an evidential requirement for comparative confidence is somehow "circular". If our aim is to give an argument for *probabilism* as an epistemic coherence requirement for comparative confidence relations, then it would seem that assuming ($\mathcal{R}_{\succeq}$) is *question-begging*.

The aim of the present monograph is not so much to provide "arguments for" particular coherence requirements (in the usual sense). Rather, the aim is to put forward a simple and strong *package* of epistemic principles, and to explain how coherence requirements — for various types of judgments — can be seen as emerging from this package in a unified way. Our package has three main components.

- (P_1) *Alethic* (or *Accuracy*) norms (and requirements). These are *extensional* (*viz.*, determined by the *truth-values* of the contents of the judgments in question) norms (and requirements). *E.g.*, (TB), (CB), (AV$_{\succeq}$), (PV$_{\succeq}$).
- (P_2) *Evidential* norms (and requirements). These are *probabilistic* (*viz.*, suitable varieties of *probabilistic representability*). *E.g.*, ($\mathcal{R}$) and ($\mathcal{R}_{\succeq}$).
- (P_3) *Coherence requirements*, which include (a) *minimization of expected inaccuracy* (relative to some evidential probability function Pr) principles, and (b) *accuracy-dominance avoidance* principles which are entailed by them. In all of our applications, the accuracy dominance avoidance principles [*e.g.*, (WADA) and (WADA$_{\succeq}$)] are offered as coherence requirements. And, in all applications of our framework, *evidential propriety* of the inaccuracy measure is essential for grounding these requirements. In the case of full belief, (P_3) is established by the proof of Theorem 1.2 (see Appendix B), and in the case of comparative confidence, (P_3) is established by the proof of Theorem 3.6 (see Appendix C).

You may have noticed that both of our "big picture" figures of norms and requirements so far (*i.e.*, Figures 1.2 and 3.4) share a common "tree" structure. Abstracting away from the details, we can now see that this "tree" always has (P_1) and (P_2) on

[93] In the context of numerical credence, the ordering assumptions are simply *baked-in via* the presupposition that numerical credence functions are *real-valued* (and that the relation $\geq$ on the real numbers is — *a fortiori* — a total order). So, we don't view the fact that we cannot offer a (qualitative) rationale for the ordering assumptions as a major shortcoming of our approach. In this sense, it seems that — from an epistemic point of view — assumptions like transitivity must be *built-in* from the outset (perhaps, as constitutive aspects of concepts describable *via* an "at least as..." locution).

its branches, and (P_3) on its its trunk (and the trunk of this “tree” is also where the epistemic rational requirements reside). Figure 4.1 gives a generic “Big Picture” representation of the logical structure of (P_1)–(P_3). Part III will have a similarly “tree” shaped “big picture” diagram (of norms and requirements for credence).

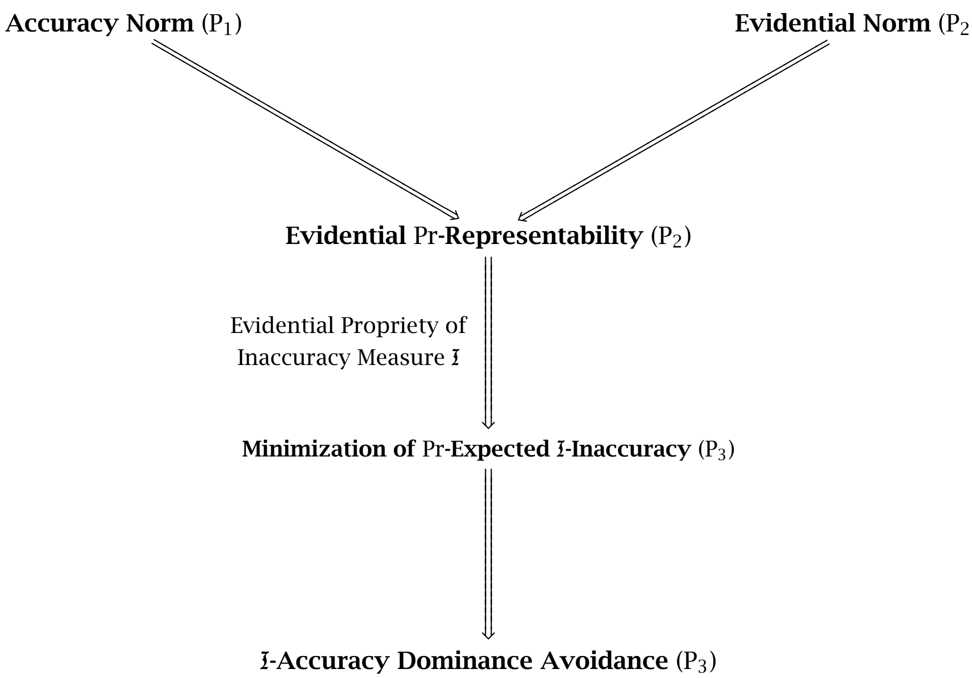

FIGURE 4.1. Generic “Big Picture” Diagram of our Package (P_1)–(P_3)

From this perspective, it is a *basic assumption* of this monograph [(P_2)] that “(numerical) probabilities reflect evidence.” This is not something that we are aiming to provide some sort of (direct or independent) “argument for” (in the traditional “arguments for probabilism” sense). Although, we will discuss this possibility in Part III below. Perhaps a better way to view our present methodology is in the spirit of a “best explication” (Carnap, 1962, Chapter 1) or a “best systems analysis” (Lewis, 1983, 1994) of formal epistemic coherence requirements.

We maintain that (P_1)–(P_3) can undergird the best *explication* of (or a “best systems analysis” of) formal epistemic coherence. Our package has many theoretical virtues, including simplicity, strength, fit, *etc.* And, we think that the quality of the package *as a whole* serves to (indirectly) lend some credence to each of its parts.

If someone were to offer a *better* explication of formal epistemic coherence (or a “better system” of requirements of epistemic rationality), then that would pose a problem for us. But, “circularity” (as such) is not a pressing worry for us. We’ll be satisfied if we manage to systematize the salient body of facts (*i.e.*, what we take to be the salient data) regarding epistemic rationality in an explanatory and illuminating way. This methodological stance will play an even more prominent role in the final part of the book (on numerical credence), to which we now turn.

Part III

Numerical Credence[94]

[94]This part of the book (especially the Negative Phase) draws on joint work with Kenny Easwaran.

CHAPTER 5

Numerical Credence: Positive Phase

So the last will be first, and the first will be last.

Matthew 20:16

1. Numerical Credence and its Formal Representation

In a historical sense, this book has been written *backwards.* The applications of the present framework to full belief and comparative confidence are very recent developments (of the past five years or so). The framework was first applied (although, as we'll see, not with exactly the same interpretation and/or methodology) to the case of numerical credence. Specifically, Roger Rosenkrantz (1981) and James Joyce (1998) were the first to give accuracy-based arguments for probabilism as an epistemic coherence requirement for numerical credence. And, Rosenkrantz and Joyce were both inspired by an elegant theorem of de Finetti (1970). We will discuss de Finetti's watershed theorem in the next section. But, first, a few preliminaries regarding the formal representation of numerical credence.

That agents (generally) believe (or disbelieve) propositions and that they (generally) make comparisons between propositions regarding their relative plausibility can hardly be questioned (of course, there are occasions on which agents ought to suspend judgment or withhold comparisons of relative plausibility). But, the claim that agents (generally) have precise, numerical degrees of confidence in propositions can seem even more dubious. At least, Keynes (1921, pp. 29–30) had doubts.

> ...no exercise of the practical judgment is possible, by which a numerical value can actually be given to the probability of every argument. So far from our being able to measure them, it is not even clear that we are always able to place them in an order of magnitude. ...The doubt, in view of these facts, whether any two probabilities are in every case even theoretically capable of comparison in terms of numbers, has not, however, received serious consideration. There seems to me to be exceedingly strong reasons for entertaining the doubt.

Nonetheless, it does seem clear that there are *some* agendas of propositions regarding which *some* agents *do* (rationally) assign precise numerical credences. In keeping with the first two parts of this book, we will focus entirely on finite agendas of propositions to which the agents we are evaluating do assign precise numerical degrees of confidence. As it happens, all of the arguments in the literature (at least, all the arguments that we will be discussing) which aim to ground coherence requirements for numerical credence make both of these assumptions of *finiteness* and *opinionation* of the judgment sets being evaluated.

That is, in our usual style, we will be considering agents S who have made numerical assignments of degree of confidence (credence) to each member of some finite agenda $\mathcal{A}$ of classical, possible-worlds propositions (drawn from some finite Boolean algebra $\mathcal{B}$). We will use the notation $b(p)$ to denote the credence S assigns to proposition p. And, we will (again, as per usual) assume our agents are logically omniscient, in the sense that they will always assign the same credences to logically equivalent propositions. For simplicity, we will also assume that our agents assign credences on the unit interval ($b(p) \in [0,1]$). That is, we will assume that there is a maximal credence (which, by convention, we'll assume to be the real number 1) and a minimal credence (which, by convention, we'll assume to be the real number 0). Finally, we will assume that our agents assign maximal credence to all tautologies [$b(\top) = 1$] and minimal credence to all contradictions [$b(\bot) = 0$]. We could provide arguments for these last two claims, but those arguments would not be very interesting. Most of the (interesting) action will involve how agents assign numerical credences to contingent claims (and logical combinations of them). So, we will (again, just for simplicity) simply pack these assumptions about $\top$ and $\bot$ into our assumption that the agents we're evaluating are logically omniscient. With that (simplified) background in place, we're ready to discuss de Finetti's Theorem.

2. The Watershed: de Finetti's Theorem

de Finetti's theorem shows that probabilism — as a coherence requirement for numerical credences — follows from a certain (formal) dominance principle. Interestingly, de Finetti did not interpret this dominance principle as an *accuracy* dominance principle. This is because de Finetti was a *radical subjectivist* (and a *pragmatist*) about probability. He did not believe that it even made sense to talk about credences as being "accurate" or "inaccurate". In this sense, de Finetti was working in the tradition of Ramsey, Savage, and others who (as we have already mentioned) have offered *pragmatic* arguments for probabilism. As a result, de Finetti did not speak of "measures of inaccuracy" or "measures of distance from vindication". Rather, he used the locutions "scoring rule" and "loss function" to talk about what we will (hereafter) be referring to as measures of inaccuracy (or measures of distance from vindication) for credence functions.[95] Be that as it may, it was de Finetti's *theorem* that inspired these later interpretations and approaches.

We will only present the simplest instance of de Finetti's theorem, because it has such an elegant geometrical interpretation (and visual representation). Consider the simple agenda $\mathcal{A} = \{P, \neg P\}$, for some contingent claim P. Given our simplified setup, an agent S's credence function $b(\cdot)$ will be *probabilistic* — relative to agenda $\mathcal{A}$ — just in case it satisfies the following additivity constraint

$$b(P) + b(\neg P) = 1.$$

Next, let's think about how we might "score" a credence function (on $\mathcal{A}$), in terms of its "*distance from the truth* (or *inaccuracy*) in a possible world". In this toy case, there are only two relevant possible worlds: w_1 in which P is true, and w_2 in which P is false (*i.e.*, $\neg P$ is true). If we use the number 1 to "numerically represent"

[95]Moreover, the original use to which de Finetti (1970) and Brier (1950) put their favorite "scoring rule" (the Brier score, to be discussed below) was for the *elicitation* of credences from agents who were *assumed to satisfy probabilism*. That is, the original purpose of "scoring rules" was not to *argue for* probabilism, but to elicit credences from agents whose credences *are* (already) *probabilistic*.

the truth-value True (at a world) and the number 0 to "numerically represent" the truth-value False (at a world), then we can "score" a credence function b using a *scoring rule* which is some function of (i) the values b assigns to P and $\neg P$, and (ii) the "numerical truth-values" of P and $\neg P$ at the two relevant worlds w_1 and w_2.

de Finetti used what is called the *Brier score* (of a credence function b, at a world w), which is just the (squared) *Euclidean distance* — *i.e.*, the sum of the squared differences between credences and (the 0/1 numerical correlates of) truth values. Relative to the trivial agenda $\mathcal{A}$, the Brier score of b is defined as follows.

$$\text{The Brier score of } b \text{ in } w_1 \stackrel{\text{def}}{=} (0 - b(P))^2 + (1 - b(\neg P))^2.$$

$$\text{The Brier score of } b \text{ in } w_2 \stackrel{\text{def}}{=} (1 - b(P))^2 + (0 - b(\neg P))^2.$$

From our perspective (which we will explain in more detail in subsequent sections of this chapter), the idea behind all such scoring rules is that "distance from truth" or "inaccuracy" of a credence function b (on agenda $\mathcal{A}$ and at world w) is measured in terms of b's "distance (at w) from the numerical truth-values of the propositions in $\mathcal{A}$". In this simple case, we can think of credence functions as 2-dimensional (real) vectors in an $\langle x, y\rangle$ space that represents possible numerical values on $[0, 1]$ which may be assigned to P and $\neg P$, respectively, by some credence function. And, we can think of the two possible worlds (w_1 and w_2) as the two unit vectors $\langle 1, 0\rangle$ and $\langle 0, 1\rangle$, respectively.[96] From this accuracy-theoretic perspective, the Brier score of b (at w) is just the (squared) *Euclidean distance* of b from w's world-vector. There is a rather vast literature on alternative "scoring rules" (or alternative measures of inaccuracy for credence functions). We will return to the controversies regarding the proper measurement of the "accuracy" of a credence function below. For now, we will just present the simplest instance of de Finetti's theorem.

> **Theorem** (de Finetti 1970). S's credence function b (on $\mathcal{A}$) is *not* probabilistic (*i.e.*, S's b *violates* the additivity constraint above) *if and only if* ($\Leftrightarrow$) there exists another credence function b' *which has lower Brier score in every possible world* (*i.e.*, b' is *closer to the numerical truth-values* of $P, \neg P$ than b is *in every possible world*, as measured *via* Euclidean distance).

The x-axes in the plots in FIGURE 5.1 represent $b(P)$ and the y-axes represent $b(\neg P)$. The diagonal lines in the plots represent the set of all of *probabilistic* credence functions b such that $b(P) + b(\neg P) = 1$. The *only if* direction ($\Rightarrow$) of de Finetti's theorem is illustrated in the plot on the left side of FIGURE 5.1. The dot (at approximately $b(P) = b(\neg P) = 1/3$) represents a non-probabilistic credence function b. The two curves drawn through the dot represent the sets of credence functions that are the same (Euclidean) distance as b from w_1 and w_2, respectively. The credence functions in the shaded region (which will be non-empty, so long as b is non-probabilistic) are *guaranteed to be closer* (in Euclidean distance) *to the truth-values* of $P, \neg P$ *in both possible worlds*. The *if* direction ($\Leftarrow$) of de Finetti's theorem is illustrated in the plot on the right side of FIGURE 5.1. This time, the dot

[96]Of course, de Finetti would not have interpreted things in this accuracy-theoretic way. But, since we are not interested in pragmatic arguments for coherence requirements, we will not delve into de Finetti's favored interpretation, which would involve viewing the Brier score (and other "scoring rules") as some sort of "pragmatic loss function" (which could measure, *e.g.*, how much money an agent is bound to lose in a given betting scenario and relative to a given possible world).

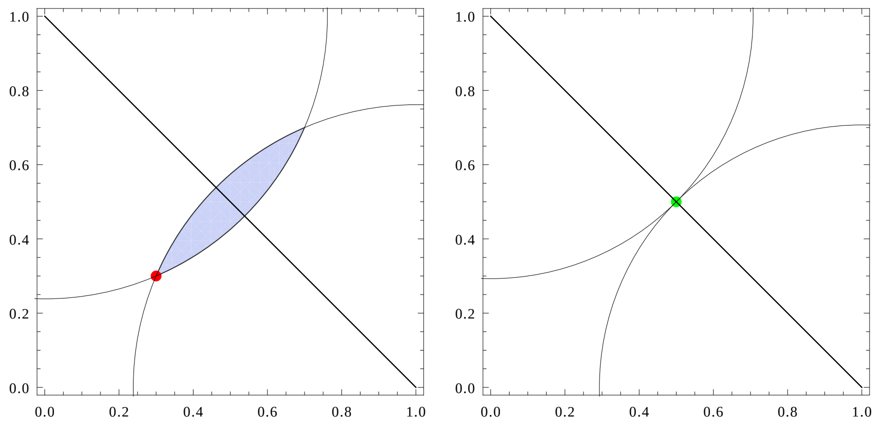

FIGURE 5.1. Geometrical proof of the two directions of (the simplest instance of) de Finetti's theorem

represents a *probabilistic* credence function b. The two curves drawn through the dot represent the sets of credence functions that are the same (Euclidean) distance as b from w_1 and w_2, respectively. This time, the curves are *tangent*, which means *there is no* credence function b' that is closer (in Euclidean distance, *come what may*) to the truth than b is. This simple geometrical argument (regarding credences on the 2-element partition $\mathcal{A}$) can be generalized to arbitrary finite agendas (and arbitrary finite Boolean algebras) of propositions (de Finetti, 1970).

3. Joyce's Re-Interpretation and Generalization(s) of Rosenkrantz's de Finetti

Roger Rosenkrantz (1981, §2.2) was the first to re-interpret de Finetti's argument in accuracy-theoretic terms. Rosenkrantz pointed out that if we interpret the Brier score as a measure of a credence function's *accuracy*, then we can view de Finetti's theorem as providing an *accuracy-dominance* argument in favor of probabilism. That, more or less, is the gloss we put on de Finetti's theorem in the previous section. In the last fifteen years or so, James Joyce (1998, 2009) has taken this accuracy-theoretic approach to grounding probabilism much farther. In this section, we will try to situate Joyce's argument(s) for probabilism within our present framework.[97] First, we'll situate Rosenkrantz's argument.

As always, these types of arguments can be seen (from our current perspective) as involving the completion of "The Three Steps". In the case of Rosenkrantz's argument(s) for numerical probabilism, these three steps are as follows.

> **Step 1**: An individual credal judgment $\ulcorner b(p) = r \urcorner$ is *inaccurate* at world w just in case r differs from the value assigned by the indicator function $v_w(p)$, which will be 1 if p is true at w and 0 if p is false at w.
>
> **Step 2**: The measure the *degree of inaccuracy* of an individual credal judgment $\ulcorner b(p) = r \urcorner$ is the (squared) *Euclidean distance* (*a.k.a.*, the Brier

[97]For a much more detailed and comprehensive survey of Joyce-style arguments for probabilism (and other norms and requirements for credences), see (Pettigrew, 2011, 2013b).

score) between r and $v_w(p)$. That is, $\mathfrak{i}_b(b(p), w) = (b(p) - v_w(p))^2$. And, the total inaccuracy of a credence function b on an agenda $\mathcal{A} = \{p_1, \ldots, p_m\}$ at a possible world w is given by the sum of these point-wise $\mathfrak{i}$-scores on $\mathcal{A}$ at w. That is, $\mathcal{I}_b(b, w) = \sum_i \mathfrak{i}_b(b(p_i), w)$.

Step 3: The *fundamental epistemic principle* is *weak accuracy dominance avoidance* for numerical credences (WADA$_b$) — b shouldn't be weakly dominated (by any b'), according to $\mathcal{I}_b(\cdot, w)$. Formally, there should not exist an alternative numerical credence function b' such that:

(i) $(\forall w)[\mathcal{I}_b(b', w) \le \mathcal{I}_b(b, w)]$, and

(ii) $(\exists w)[\mathcal{I}_b(b', w) < \mathcal{I}_b(b, w)]$.[98]

Joyce (1998, 2009) has offered various generalizations of Rosenkrantz's argument. In his first paper, Joyce (1998) offered a set of axioms for "the overall gradational inaccuracy of b on agenda $\mathcal{A}$ at possible world w", and he showed that the above argument would go through for *any* inaccuracy measure satisfying those axioms (holding Steps 1 and 3 fixed). In his later paper, Joyce (2009) proved a result that is in one sense more general and in one sense less general than his earlier theorem. The later theorem is more general in that it applies to a broader class of measures of inaccuracy, and it is less general in that it applies only to *partitions*, and not to *arbitrary agendas* $\mathcal{A}$ of propositions. I will not discuss Joyce's later argument here, since I want to focus on arguments that will work for arbitrary judgment sets (and not just sets of judgments regarding the elements of a partition).

There have been various aspects of Joyce's earlier argument that have been the source of controversy. Some have questioned his explication of the inaccuracy of an individual credal judgment $\ulcorner b(p) = r \urcorner$ at a possible world w in Step 1 (Hájek, 2010). Some have questioned his choice of inaccuracy measure $\mathcal{I}_b$ in Step 2 (Maher, 2002). And, some have questioned his choice of fundamental epistemic principle (Caie, 2013; Fitelson, 2012; Easwaran and Fitelson, 2012). We will address each of these worries in the Negative Phase, below. But, first, I would like to discuss a more recent argument for numerical probabilism, which is a different sort of generalization of Rosenkrantz's accuracy-theoretic take on de Finetti.

4. The Argument From Propriety

Predd et al. (2009) have given a particularly probative (and general) version of the sort of argument for probabilism we've been discussing. They prove that, if we hold Steps 1 and 3 fixed in the above description of Rosenkrantz's argument, then Step 2 can be replaced by the following, while still yielding *probabilism* as a coherence requirement for numerical confidence.

Step 2 (Predd et al., 2009). The measure the *degree of inaccuracy* of an individual credal judgment $\ulcorner b(p) = r \urcorner$ at a possible world w is *any continuous proper scoring rule* $\mathfrak{i}(b(p), w)$. And, the total inaccuracy of a credence function b on an agenda $\mathcal{A} = \{p_1, \ldots, p_m\}$ at a possible world

[98]The requirement of *strict* accuracy dominance avoidance (SADA$_b$) has sometimes been adopted as the fundamental epistemic principle in these arguments. But, for most scoring rules people have actually used (*i.e.*, for any continuous and *strictly* proper measure of credal inaccuracy), (WADA$_b$) and (SADA$_b$) are *equivalent* (see *fn.* 34). In keeping with first two parts of the book, we will use (WADA$_b$) here. But, see *fn.* 79 for more on the relationship between strict propriety and (SADA$_b$).

w is given by the sum of these point-wise $\mathfrak{i}$-scores on $\mathcal{A}$ at w. That is, $\mathcal{I}(b, w) = \sum_i \mathfrak{i}(b(p_i), w)$.

Where a *proper scoring rule* is simply a point-wise measure of inaccuracy that satisfies the following definition.

> **Definition** (propriety). A point-wise measure $\mathfrak{i}(b(p), w)$ of the *degree of inaccuracy* of an individual credal judgment $\ulcorner b(p) = r \urcorner$ at a possible world w is said to be **proper** just in case every *probabilistic* credence function b minimizes *expected* $\mathfrak{i}$-inaccuracy, *according to itself.*

In other words, $\mathfrak{i}$ is proper iff whenever we use a *probabilistic* credence function b to calculate *expected* $\mathfrak{i}$-inaccuracy, no other credence function b' will have a lower expected $\mathfrak{i}$-inaccuracy than b itself. The expected $\mathfrak{i}$-inaccuracy of an individual credal judgment b' — according to a *probabilistic* credence function b — is:

$$\sum_w b \cdot \mathfrak{i}(b'(p), w)$$

If this quantity is always minimized when $b' = b$, then $\mathfrak{i}$ is a proper scoring rule.

The (squared) Euclidean distance (*i.e.*, the Brier score) is but one of many (continuous) proper scoring rules (*viz.*, point-wise measures of inaccuracy). As a result, the theorem of Predd et al. (2009) is a vast generalization of the original theorem of de Finetti (and also those of Rosenkrantz and Joyce). Moreover, this theorem is exactly the theoretical device we need in order to apply our present framework to the problem of grounding epistemic coherence requirements for b.

5. Our "Probabilist-Evidentialist Spin" on the Argument from Propriety

Recall our definition of *evidential propriety* (Definition 1.3, p. 22).

> **Definition** (evidential propriety). Suppose a judgment set **J** of type $\mathfrak{J}$ is supported by the evidence. That is, suppose there exists some evidential probability function $\Pr(\cdot)$ which represents **J** (in the appropriate sense of "represents" for judgment sets of type $\mathfrak{J}$). If this is sufficient to ensure that **J** minimizes expected inaccuracy (relative to Pr), according to the measure of inaccuracy $\mathfrak{I}(\mathbf{J}, w)$, then the measure $\mathfrak{I}$ is **evidentially proper**.

In the case of full belief, the suitable notion of "evidential probabilistic representation" was given by the following condition (p. 19).

($\mathcal{R}$) There exists some probability function that probabilifies (*i.e.*, assigns probability greater than $1/2$ to) each belief in **B** and dis-probabilifies (*i.e.*, assigns probability less than $1/2$ to) each disbelief in **B**.

In the case of comparative confidence, the suitable notion of "evidential probabilistic representation" was given by the following condition (p. 60).

($\mathcal{R}_{\succeq}$) There exists some regular probability function $\Pr(\cdot)$ such that:

(i) $\Pr(p) > \Pr(q)$, for each $p, q \in \mathcal{A}$ such that $p \succ q$,

(ii) $\Pr(p) = \Pr(q)$, for each $p, q \in \mathcal{A}$ such that $p \sim q$.

The question at hand is: What is the suitable notion of "evidential probabilistic representation" in the case of numerical credence? It would seem to be this.

($\mathcal{R}_b$) There exists some probability function $\Pr(\cdot)$ such that, for each $p \in \mathcal{A}$,

$$b(p) = \Pr(p).$$

In other words, if one assumes that "numerical probabilities reflect evidence", then one is led to the evidential requirement ($\mathcal{R}_b$) that one's numerical credences should *match* some (evidential) probability function (and therefore some probability function). Of course, this would be an *even more circular* "argument for probabilism" than the "argument for (regular) probabilistic representability of $\succeq$" that we saw in Part II. As we explained above (Part II, §5), this is the wrong way to "spin" our story.

A preferable way to think about what we're up to is this. According to the present approach, the way to think about coherence requirements (for judgments of type $\mathfrak{J}$) is as (formal) wide-scope epistemic requirements that are entailed by both the alethic and the evidential norms/ideals (for judgments of type $\mathfrak{J}$). In the case of full belief, we argued (on these grounds) that accuracy-dominance avoidance requirements like (WADA) are good candidates for epistemic coherence requirements (and better candidates than deductive consistency, *viz.*, PV, which is *too demanding*, in light of Preface cases and the like). Analogously, in the case of comparative confidence, we argued that accuracy-dominance avoidance requirements like (WADA$_{\succeq}$) are good candidates for epistemic coherence requirements (and better candidates than the analogue of deductive consistency for comparative confidence judgments, PV$_{\succeq}$, which is perhaps even more implausible than PV). In both of those cases, the theoretical key to our argument was the adoption of an *evidentially proper* measure $\mathfrak{I}$ of doxastic inaccuracy. The evidential propriety of our measures $\mathfrak{I}$ is what guaranteed that our accuracy-dominance avoidance requirements followed from both our alethic and our evidential norms/ideals (see Figure 4.1 on p. 71). That brings us to the case of numerical credence.

In the case of numerical credence, the following analogue of deductive consistency is *utterly implausible* as a requirement of epistemic rationality.

(PV$_b$) There should be some possible world w at which the credence function b is *vindicated* (*i.e.*, for some w, b is identical to w's indicator function v_w).

This is *utterly implausible* as a requirement of epistemic rationality, because it would require all epistemically rational credence functions to be *extremal*.[99] On the other hand, accuracy-dominance avoidance requirements like (WADA$_b$) seem to be much better candidate requirements of epistemic rationality for credences.[100]

The theorem of Predd et al. (2009) is the theoretical key that allows us to unlock an argument in favor of (WADA$_b$) that is continuous with the arguments we have given for accuracy-dominance avoidance requirements in Parts I and II. It is trivially true that if a measure of credal inaccuracy is proper (in the traditional sense of that term used in the previous section), then it is *evidentially* proper (in our sense). After all, every evidential probability function is a probability function.

[99]Some have used this fact as the main premise in an argument against Joyce's definition of the (qualitative) inaccuracy of individual credal judgments in Step 1 (Hájek, 2010). We will return to this criticism of the present approach in the Negative Phase.

[100]But see the Negative Phase for some potential worries regarding Step 3.

So, if an inaccuracy measure $\mathfrak{I}$ is proper in the sense of Predd et al. (2009), then it is *evidentially* proper in our sense. And, now, we have an argument for (WADA$_b$) as an epistemic rational requirement for numerical credence that is analogous to the arguments offered in Parts I and II for (WADA) and (WADA$_{\succeq}$). The analogy is not perfect, since, as it happens, in the case of numerical credence, (WADA$_b$) is *equivalent* to probabilism [*viz.*, ($\mathcal{R}_b$)]. In the case of full belief, we saw that the analogue of probabilism ($\mathcal{R}$) was *strictly stronger* than (WADA). Similarly, in the case of comparative confidence, we saw that the analogue of probabilism ($\mathcal{R}_{\succeq}$) was *strictly stronger* than (WADA$_{\succeq}$), which only implies the (full) representability of comparative confidence orderings $\succeq$ by some (Dempster-Shafer) belief function.

We have almost all the ingredients needed to construct our "big picture" of the norms and requirements for numerical confidence (*viz.*, credence). There are just two remaining (narrow-scope, credal) norms that need to be explicitly articulated.

(AV$_b$) S's credence function b should be *actually* vindicated (*i.e.*, it should match the indicator function for the *actual* world $v_{w_@}$).

(E$_b$) S's credence function b should match the *evidential* probability function (*i.e.*, the probability function which gauges the evidential support relations in the context in which the credal judgments are formed).

$$\begin{tikzcd}
(\mathrm{AV}_b) \arrow[d, Rightarrow] & & (\mathrm{E}_b) \arrow[ddl, Rightarrow] \\
(\mathrm{PV}_b) \arrow[dr, Rightarrow] & & \\
& (\mathcal{R}_b) \arrow[d, Leftrightarrow] & \\
& (\mathrm{WADA}_b) \arrow[d, Rightarrow] & \\
& (\mathrm{SADA}_b) &
\end{tikzcd}$$

FIGURE 5.2. Logical relations between norms and requirements for numerical credence

As far as we are concerned, the fact that one[101] of these logical arrows is a *double*-arrow (whereas, in the first two parts of the book, the analogous arrow was *merely a single*-arrow) does not render the present "spin" on the argument from propriety

[101] Specifically, the arrow connecting ($\mathcal{R}_b$) & (WADA$_b$). Here, we're assuming only that our scoring rule is continuous, additive, and proper. If we assume that the scoring rule is continuous, additive, and *strictly* proper (replace "minimizes expected inaccuracy" with "*uniquely* minimizes expected inaccuracy" in the definition of propriety), then the arrow connecting (WADA$_b$) & (SADA$_b$) would also become a double-arrow (Schervish et al., 2009; Pettigrew, 2011). Most credal inaccuracy measures used in the literature are *strictly* proper; so this arrow is usually a double-arrow (*fns.* 34 and 79).

“circular”. In fact, one could turn this charge of “circularity” (*qua* “argument for probabilism”) on its head, and parlay it into a “argument for” *propriety*.

Those who view the argument from propriety as an *argument for probabilism* have felt the need to give arguments in favor of propriety as a constraint on measures of doxastic inaccuracy (see Negative Phase). If you like, you can use our “spin” on these arguments as an “argument for propriety”. That argument would involve explaining that propriety is *precisely* the property an inaccuracy measure *must* have *in order to align the evidential and alethic norms/ideals and requirements for credences* — *i.e.*, in order to ensure that (WADA$_b$) follows from both the alethic and the evidential norms/ideals for numerical credence. This could be seen as an “argument for propriety”, but we would prefer to interpret what we are doing in a more holistic way.

As we explained above in section §5 of Part II, we would prefer to view the present approach as the “best explication” (or “best systems analysis”) of formal, epistemic coherence requirements (in general). In that sense, we don’t really like to think of what we’re doing as giving “arguments for” certain *parts* of the present framework. Rather, we prefer to judge the framework as a whole, in terms of the quality of the explications of coherence requirements that it yields.

CHAPTER 6

Numerical Credence: Negative Phase

> Probability does not exist.
>
> Bruno de Finetti

1. Worries about Step 1: Credal Inaccuracy

Joyce assumes that there is a unique, *vindicated* credence function $\mathring{b}_w$ in each possible world — the indicator (or valuation) function (v_w) which assigns 1 to p if p is true at w and 0 to p if p is false at w. Basically, this approach just associates the number 1 with the truth-value T and 0 with the truth value F, and then it treats $b(p)$'s distance from p's "numerical truth-value" at w as $b(p)$'s inaccuracy at w. Thus, b's overall distance from $\mathring{b}_w(p)$ — on the salient agenda $\mathcal{A}$ — is then identified with the overall doxastic inaccuracy of b on $\mathcal{A}$ at w. This has the consequence that the alethically ideal or perfectly accurate credence function (or even a *possibly* perfectly accurate credence function) is *extremal*.

Those who defend *regularity* as a rational requirement for credences (Jeffrey, 1992) may balk at this, since a defender of regularity might well think that the "ideal" credence to assign to a contingent claim p cannot be extremal, since one should (where this "should" is, presumably, *evidential* in nature — *fn.* 13) only assign extremal credences to logical truths or falsehoods.

Moreover, one might worry about this approach to credal vindication on the grounds that vindication for credences should be more closely analogous to vindication for belief (Hájek, 2010). As we discussed in Part I, the truth norm for belief (TB) [*viz.*, actual vindication (AV)] has many advocates. But, or so the objector presumes, nobody would defend a "norm of actual vindication" for credences (AV_b). After all, if Joyce is right about vindication, then it would seem that rational credences are *almost never (or, assuming regularity, never) vindicated.* This leads Hájek (2010) to seek an alternative (narrow-scope) alethic norm for credences. He proposes an alternative according to which the vindicated credence function $\mathring{b}_w$ is *the chance function* at w (Ch_w). While it is true that the chance function can handle problems involving the *extremality* of $\mathring{b}_w$, we think this alternative is incorrect. Indeed, we think that — *if* there is a vindicated credence function $\mathring{b}_w$ (we'll give some additional reasons to be skeptical about this, below) — the vindicated credence function must be the indicator function.

There are several ways to see the inevitability of Joyce's explication of $\mathring{b}_w$ (if there is such an explicandum). First, as we explained in Part II, there is a plausible alethic sense of vindication for *comparative confidence* relations $\succeq$ (*i.e.*, ranking all truths at w strictly above all falsehoods at w). And, if this is right, then how could it be a category mistake to suppose there is an analogous alethic ideal for numerical credence? After all, many have maintained that numerical credences b should serve as *representers* of comparative confidence relations $\succeq_b$. On this view, if b were *identical* to the indicator function, then the comparative confidence relation $\succeq_b$ given rise to by b would be accurate in the sense of Part II.[102] Indeed, if $\mathring{b}_w$ is taken to be *any non*-extremal function, then this (general) correspondence between vindication (or non-inaccuracy) for numerical credence functions b and the comparative confidence relations $\succeq_b$ they represent would break down.

We think the main problem with taking the chance function (or anything extensionally less informative than the indicator function) as the vindicator of credence functions is that this move seems to conflate alethic and evidential norms for judgment. The Principal Principle (Lewis, 1981) is an *evidential* norm for credence, which says (roughly) *if* your total evidence consists of knowledge of the chance of p, *then* your credence in p should match the chance function (*viz.*, $b(p) = \mathrm{Ch}_w(p)$). But, of course, if your total evidence includes *inadmissible evidence* about p, then that evidence will trump the evidence you have regarding the chance of p. Specifically, it seems plausible to suggest that if you know the value the indicator function assigns to p at w ($v_w(p)$), then you should assign *that* credence to p (*viz.*, $b(p) = v_w(p)$).[103] These considerations also seem to support the Joycean explication of $\mathring{b}_w$.

Finally, we should address the worry that it is odd to defend an alethic "norm" that is *almost never* (or, if regularity is a rational requirement, *never*) satisfied by rational agents. In contrast, the truth norm for belief (TB) is often satisfied by rational agents, even if (as we argued in Part I) it is not a rational requirement. This worry does not seem decisive. The fact that some *ideal standard* is almost never (or never) achieved by rational agents does not necessarily imply that the standard in question is inapt (*qua ideal* standard). For instance, consider the epistemic ideal of *omniscience*. While no rational agent is (or could be) omniscient, this does not imply that omniscient agents are not *epistemically ideal*. And, something similar can be said for Joyce's explication of $\mathring{b}_w$. Just because no rational agent (let us suppose) is (exactly) vindicated in Joyce's sense, this does not imply that vindicated credence functions (in Joyce's sense) are not *alethically ideal/perfect*. Moreover, because we are mainly interested in *how close* an agent's credence function is to alethic perfection, we needn't be too bothered by the fact that we can never achieve (complete) alethic perfection. If one agent knows more than another, then they may be said to be epistemically *better* because they are *closer* to the ideal of omniscience. Similarly, if one agent's credence function is closer to Joyce's $\mathring{b}_w$, then they may be said to be *closer* to the alethic ideal of vindication.[104] Finally, recall comparative confidence judgments of the form $P \sim \neg P$. Such judgments will *inevitably be inaccurate*, since they entail that the relation $\succeq$ fails to fully separate the truths and the falsehoods. Nonetheless, such judgments can be supported by

[102] In fact, the resulting Joycean notion of an "accurate comparative confidence relation $\succeq$" is *strictly stronger* than ours — see *fn.* 74.

[103] Although, one might worry that this principle is only plausible if you know $v_w(p)$ *with certainty*. That is, one might think that Joyce's explication of vindication for degree of confidence might be better suited to serving as a vindicator for *certainties* (and not *uncertain*, but extremal, numerical credences). I will not discuss this further, but there may be a serious worry lurking here.

[104] This explains why Joyce's $\mathring{b}_w$ is sometimes glossed as "the credence function that the alethically ideal (or omniscient) agent would have (at w)."

the evidence (*e.g.*, if the evidence regarding P and $\neg P$ is *balanced* in the right way). In the credal case, this phenomenon is *more widespread* (indeed, it's *nearly universal* if regularity holds). But, this (in and of itself) doesn't seem to be a reason to reject Joyce's explication of $\mathring{b}_w$.

Ultimately, we think that (*if* there is an explicandum $\mathring{b}_w$ — more on this in subsequent sections) Joyce's explication of $\mathring{b}_w$ is best. In any event, even if this turns out to be wrong, we could revert to our account of coherence requirements for *comparative confidence*, which seems to involve a notion of "vindication" or "accuracy" that avoids many of the worries people have had about Joyce's $\mathring{b}_w$. Of course, that approach wouldn't — using ($\text{WADA}_{\succeq}$) *alone* — get us all the way to probabilism ($\mathfrak{C}_4$), but it would at least imply that our comparative confidence orderings should be representable by some Dempster-Shafer belief function ($\mathfrak{C}_1$). Anyhow, we'll have more to say about some of these issues below. Meanwhile, let's move on to worries that have been voiced regarding Joyce's Step 2.

2. Worries about Step 2: The Scoring Rule Controversy

This is perhaps the most controversial Step of the Joycean approach. There have been many objections to Joyce's argument that take issue with some aspect or other of this second step (Maher, 2002; Bronfman, ms; Hájek, 2008; Pettigrew, 2011). I will not delve into all of these objections. Instead, I will focus on a particular line of attack that was initiated by Maher (2002). According to this line, we are owed an *argument* for the claim that we should use some proper score (like the Brier score), as opposed to some improper score. Specifically, Maher (2002) asks why we *shouldn't* use the following *improper* measure of the inaccuracy of individual numerical credence judgments — the L_1-norm (Deza and Deza, 2009).

$$L_1(b(p), w) \stackrel{\text{def}}{=} \left| b(p) - \mathring{b}_w(p) \right|$$

Joyce (2009) gives a compelling response to this query of Maher's. He points out that L_1 has some undesirable behavior, when combined with (WADA_b) and/or (SADA_b). Specifically, according to the additive measure of the total inaccuracy (of b at w) generated by L_1, some credal states that seem to be supported by the evidence will be *ruled out* by accuracy-dominance considerations.

Let P_i be the claim that a 3-sided die which has been tossed came up "#i". Suppose S has the credal set $b = \langle 1/3, 1/3, 1/3 \rangle$. And, suppose S *knows only that the die is fair* (*i.e.*, S has *no other P_i-relevant evidence*). It seems (especially from our probabilist-evidentialist perspective) that b should not be ruled out as epistemically irrational. Unfortunately, $b' = \langle 0, 0, 0 \rangle$ *strictly L_1-dominates b*. That is, according to L_1, the credal state of *assigning minimal credence to all three possible outcomes P_i* is strictly more accurate than the indifferent credal assignment b in every possible world. Assuming that (SADA_b) is a rational requirement (given a suitable measure of credal inaccuracy), if we want our coherence requirements to be sensitive to evidential norms for credences (as *we* do), then this would seem to be a telling argument against L_1. This is similar to the argument we gave against naïve, mistake-counting measures of inaccuracy for comparative confidence relations (*e.g.*, Δ and the Kemeny measure). We are sympathetic to this Joycean strategy for arguing against the use of certain measures of doxastic inaccuracy.[105]

More generally, the potential for conflict between alethic and evidential requirements is one of the central recurring themes of this book. Preface cases reveal that deductive consistency can come into conflict with evidential requirements for full belief. And, examples in which measures of doxastic inaccuracy that cause (*via* accuracy dominance) some evidentially supported states to be *ruled out* reveal that such measures are inappropriate (*viz.*, evidentially improper). In the next section, we'll discuss some worries about accuracy dominance avoidance requirements (and how they interact with other requirements, especially evidential ones).

3. Worries about Step 3: Three Worries About Accuracy Dominance

3.1. An "Evidentialist" Worry About Joyce's Argument[106]. Suppose an agent S is trying to figure out which candidate credence functions to *rule out*. In this process, S might appeal to various sorts of norms and considerations. For instance, S might want to rule out credence functions that are susceptible to a Dutch Book (DB). Moreover, S might want to avoid credence functions that violate the Principal Principle (Lewis, 1981) (PP), given her current knowledge concerning objective chance.

(PP) If S knows that the objective chance of p is less than r, and S has no inadmissible evidence regarding p, then S should not assign a credence greater than r to p.

Suppose (1) S rules out those credence functions that are susceptible to Dutch Book (DB), and then (2) S rules out those credence functions that violate (PP) given her current knowledge K concerning objective chance. This two-step procedure will rule out the same set of credence functions as the procedure which performs (2) first, and then (1). Specifically, no non-probabilistic b's will survive either two-step ruling out procedure. In this sense, *the order in which the two norms (DB) and (PP) are applied does not matter.*

Interestingly, this order-independence property seems to be *violated* by accuracy dominance avoidance norms like (WADA_b) and (SADA_b). For instance, suppose (a) S rules out those credence functions that are strictly (Brier) accuracy-dominated (SADA_b) by a credence function that is not yet ruled out, *and then* (b) S rules out

[105]Indeed, we suspect that *any* improper scoring rule (or *any* evidentially improper inaccuracy measure) will *inevitably* rule out (*via* weak dominance) some (evidentially) probabilistically representable doxastic states. This conjecture, if true, would allow us to generalize our arguments from propriety — it would allow us to replace the assumption of propriety with the (*prima facie* more plausible) assumptions of Coherent Admissibility (*i.e.*, never ruling out, *via* dominance, a probabilistically representable doxastic state), Truth-Directedness (*fn.* 30), and Continuity (Joyce, 2009). For more on this conjecture and its ramifications, see (Pettigrew, 2011, Conjecture 1).

[105]This section draws heavily on (Easwaran and Fitelson, 2012). Indeed, in this preliminary draft, the section is is largely *lifted from* that paper. Note: as we have explained above, our present arguments are not best understood as "arguments for probabilism" in the usual sense. We're *assuming* a probabilistic rendition of evidentialism. So, the examples discussed in this section are *not* problems for *us*, since we're *assuming* that "(numerical) probabilities reflect evidence". But, the arguments in this section are among the reasons we've decided to adopt this stance in the book.

those credence functions that violate (PP) given her current knowledge K concerning objective chance. No non-probabilistic credence functions b can survive *this* two-step procedure. But, if we *reverse the order* — that is, if we perform (b) first *and then* (a) — then some non-probabilistic credence functions *can* survive.

Here is a simple example that illustrates this *interaction/order-effect*. Suppose S's background knowledge K contains (*exactly*) the following information about the chance of P (and *no inadmissible evidence*).

(K) The objective chance of P is at most 0.2.

We can understand the effect by looking at the diagrams in Figure 6.1. As in Figure 5.1 above, the square represents the set of credence functions taking values between 0 and 1, with x-axis representing the agent's credence in P and the y-axis representing the agent's credence function in $\neg P$. For a given credence function b (represented by the dot) the two circular arcs delineate the regions that are at least as accurate as b in worlds w_2 (the upper-left corner) and w_1 (the lower-right corner).[107] The shaded region in the left diagram then represents the set of credence functions that accuracy-dominate b. The green region in the right diagram represents the credence functions that satisfy (PP) given knowledge K.

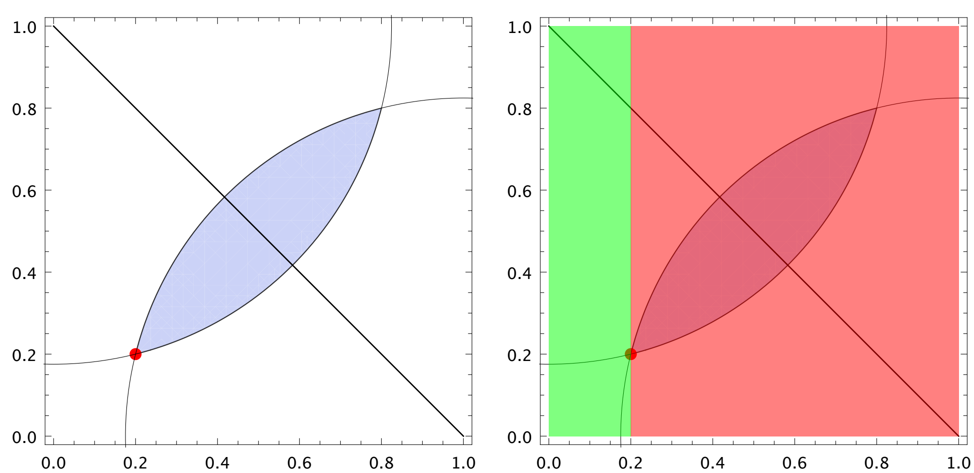

FIGURE 6.1. Visualizing the order-dependence of (SADA_b) and (PP)

Thus, if the agent first applies (a), she will rule out every credence function that is not on the diagonal line, because they have a non-empty shaded region on the left diagram. Applying (b) second, she will then rule out every remaining credence function in the red region on the right diagram, leaving just the upper-left part of the diagonal line.

However, if the agent first applies (b), she will rule out every credence function in the red region on the right diagram. When she then applies (a), what happens

[107]Here, we're assuming the Brier score (Euclidean distance) as our inaccuracy measure. This is inessential to the present worry. We could run an analogous argument for any proper scoring rule. The Brier score is convenient here, since it has such a nice geometrical interpretation.

will be different. For a credence function like b (which survives application of (b), since it is part of the green region), (a) will not say anything, because the only b' that dominate it have already been ruled out by (b). As Theorem 6.1 below will show, the functions that survive this order of application will be the ones on the upper-left part of the diagonal line, *and also all the ones on the border of the green and red regions that are below the diagonal line.*

In the toy case where the algebra consists only of four propositions, and credences in $\top$ and $\bot$ are fixed at 1 and 0 respectively, the following result applies:

THEOREM 6.1. *If b' dominates b, and both take values only in $[0,1]$, then:*

either $b(P) > b'(P)$ *and* $b(\neg P) > b'(\neg P)$ or $b(P) < b'(P)$ *and* $b(\neg P) < b'(\neg P)$.

An agent with credence function b will evaluate herself as having violated a norm if she applies (a) before (b), but not if she applies (b) before (a).

In fact, our worry is much more general than this example involving (PP) suggests. Joyce's argument tacitly presupposes that — for any incoherent agent S with credence function b — some (coherent) functions b' that Brier-dominate b are always "available" as "permissible alternative credences" for S. But, there are various reasons why this may not be the case.[108] The agent could have good reasons for adopting (or sticking with) *some* of their credences. And, if they do, then the fact that some accuracy-dominating (coherent) functions b' "exist" (in an abstract mathematical sense) may not be epistemologically probative, from their current epistemic perspective. Thus, our use of the Principal Principle (PP) is merely one illustration of this more general phenomenon. Therefore, while one may object to our use of (PP) here for various reasons[109], our worry will remain pressing, provided only that the following sorts of cases are possible.

($\star$) Cases in which (a) S's credence function b is not a probability function, (b) S assigns $b(p) \in \mathcal{I}$, for some p and some interval $\mathcal{I}$, (c) S has good reason to believe (or even *knows*) that epistemic rationality (*viz.*, the evidential norm for credences) *requires* $b(p) \in \mathcal{I}$, *but* (d) all the credence functions b' that Brier-dominate S's credence function b are such that $b'(p) \notin \mathcal{I}$.

We have tried to describe a simple, toy case satisfying ($\star$) — by making use of (PP). Even if one thinks our toy (PP) example is infelicitous (see *fn.* 109), this won't be enough to make our worry go away. In order to avoid our worry completely, one would need to argue that no examples satisfying ($\star$) are possible. And, *that* is a tall order. Surely, we can imagine that an *oracle* concerning epistemic rationality has

[108] Presently, we're concerned with *evidential* reasons why such b's may be unavailable to an agent. There may also be *psychological* reasons why some "alternative b'-functions" may be unavailable, but we are bracketing that possibility here.

[109] For instance, the (PP) was originally intended Lewis (1981) to be applied to agents with "reasonable" — indeed, *probabilistic* — credence functions. In our case, the agent doesn't recognize, for instance, that K entails that the objective chance of $\neg P$ is at least 0.8. While this may be hard to imagine for such a small example, if the algebra is sufficiently large, then it becomes plausible that the agent won't recognize all the constraints on objective chance.

In addition, the (PP) was originally intended to be applied to *initial* credence functions, which would not be informed by specific bodies of empirical knowledge regarding objective chances (*e.g.*, our K above). For these reasons, our present applications of (PP) are not (strictly speaking) kosher.

Ultimately, however, our worry will remain — so long as examples satisfying ($\star$) are possible (see below). And, we see no reason to doubt that such examples exist.

informed S that $b(p) \in \mathcal{I}$ is *required* — despite the fact that all (coherent) Brier-dominating functions b' are such that $b'(p) \notin \mathcal{I}$. While such cases are fanciful, it seems to us that they are sufficient to motivate our worry.

It is interesting to note that Dutch Book arguments do *not* have this feature. As far as (DB) is concerned, it doesn't matter if you have good reasons for sticking with some of your credences. Suppose you do. Nonetheless, it remains true that if (and only if) you're incoherent, you're susceptible to Dutch Book. And, this gives you some reason (albeit a pragmatic reason) to change your *other* credences, so as to bring yourself into a coherent doxastic state. In the example(s) depicted in Figure 1, for instance, (DB) would give S/S' some reason (albeit a pragmatic one) to change their credences to a probabilistic function in the green region. So, this "order-dependence" is a *peculiarity* of "accuracy-dominance"-based approaches to probabilism.

Note that the structure of the present worry is very similar to the structure of Joyce's own argument against L_1. If someone adopts L_1, then they seem to run into conflicts between an evidential requirement (PP) and an alethic requirement (SADA$_b$). And, Joyce takes this to be a reason to reject the distance measure L_1. In the present context, the worry is that an agent could run into another sort of "conflict" between (PP) and (SADA$_b$), *even if* they adopt the Brier score $\acute{\imath}_b$ as their measure of credal inaccuracy. In this case, rejecting the distance measure doesn't seem to be an apt response. The only other options would seem to be rejecting (PP) or rejecting (SADA$_b$).

This requires some consideration of what the role of the Principal Principle is in constraining belief. If it helps determine which credence functions are available, then it seems that both of Joyce's arguments run into problems. One might instead think that violation of the Principal Principle doesn't make a credence function unavailable, but instead just represents some dimension of epistemic "badness". If this badness is different from the badness of inaccuracy, then it becomes clear that Joyce's arguments need to be modified — even if b' dominates b with respect to inaccuracy, if b has less *overall* epistemic badness, then b may still be perfectly acceptable as a credence function. Thus, Joyce's arguments would need to consider overall badness rather than just inaccuracy.

The only way to save Joyce's arguments here seems to be to say that somehow the badness of violating the Principal Principle is *already included* when one has evaluated the accuracy of a credence function. Perhaps there is some way to argue for this claim. And nothing here turns on the use of the Principal Principle in particular — if there can be any epistemic norm whose force is separate from accuracy, then the same sort of problem will arise. Joyce's argument seems to work only if *all* epistemic norms spring from accuracy.[110]

3.2. Is Joyce's Argument Language Dependent? Suppose we have two numerical quantities ϕ and ψ. These might be, for instance, the velocities (in some

[110] In recent and forthcoming work, Joyce (2013) and Pettigrew (2013a) have attempted to articulate "accuracy-first" epistemologies that exploit varieties of (alethic) epistemic value monism in order to respond to worries such as these. We prefer our approach, which simply *assumes* that "probabilities reflect evidence"; and, hence, that examples of the sort described in ($\star$) above *do not exist*. We will return to some of these questions below, when we consider "accuracy-first epistemology" and how it might reply to some of the worries we have discussed in the Negative Phase.

common units) of two objects, at some time (or some other suitable physical quantity of two objects at a time). Suppose further that we have two sets of predictions concerning the values of ϕ and ψ, which are entailed by two hypotheses H_1 and H_2, and let's denote the truth about the values of ϕ and ψ (or, if you prefer, the true hypothesis about their values) — in our standard units — as T. Let the predictions of H_1 and H_2, and the true values T of ϕ and ψ be given by the following table. [Ignore the α/β columns of the table, for now — I'll explain the significance of those columns, below.]

	ϕ	ψ	α	β
H_1	0.150	1.225	0.925	2.000
H_2	0.100	1.000	0.800	1.700
T	0.000	1.000	1.000	2.000

TABLE 6.1. Canonical example of the language dependence of the accuracy of numerical predictions

It seems clear that the predictions of H_2 are "closer to the truth T about ϕ and ψ" than the predictions of H_1 are. After all, the predicted values entailed by H_2 are strictly in between the values predicted by H_1 and the true values entailed by T. However, as Popper (1972, Appendix 2) showed [using a recipe invented by David Miller (2006)], there exist quantities α and β (as in the table) satisfying both of the following conditions.

(1) α and β are symmetrically inter-definable with respect to ϕ and ψ in the following (linear) way:

$$\begin{aligned} \alpha &= \psi - 2\phi & \qquad \beta &= 2\psi - 3\phi \\ \phi &= \beta - 2\alpha & \qquad \psi &= 2\beta - 3\alpha \end{aligned}$$

(2) The values for α and β entailed by H_2 are strictly "farther from the truth T about α and β" than the values for α and β entailed by H_1.

As Miller (1975) explains [see (Miller, 2006, Chapter 11) for a nice historical survey], there is a much more general result in the vicinity. It can be shown that for any pair of false theories H_1 and H_2 about parameters ϕ and ψ, *many* comparative relations of "closer to the truth" between H_1 and H_2 regarding ϕ and ψ can be *reversed* by looking at what the estimates provided by H_1 and H_2 for ϕ and ψ entail about quantities α and β, which are symmetrically inter-definable with respect to ϕ and ψ, *via* some (linear) inter-translation of the form:

$$\begin{aligned} \alpha &= \mathsf{a}\psi + \mathsf{b}\phi & \qquad \beta &= \mathsf{c}\psi + \mathsf{d}\phi \\ \phi &= \mathsf{a}\beta + \mathsf{b}\alpha & \qquad \psi &= \mathsf{c}\beta + \mathsf{d}\alpha \end{aligned}$$

That is, for *many* cases in which we judge that "H_2 is closer to the truth T about ϕ and ψ than H_1 is" (on many ways of comparing "closeness") there will exist some member of the above family of symmetric inter-translations such that we will judge that "H_1 is closer to the truth T about α and β than H_2 is". In this way, we can often *reverse* accuracy comparisons of quantitative theories *via* such re-descriptions of prediction problems. As such, many assessments of the accuracy of predictions are *language dependent*.

According to Joyce (1998), if we view credences (of rational agents) as *numerical estimates of truth-values of propositions*, then we can give an argument for probabilism that is based on considerations having to do with the "accuracy" of such estimates. We will focus on a simple, concrete example that illustrates a (potential) problem of language dependence.

Consider an agent S facing a very simple situation (just like the other situations we've been considering so far in Part III), involving only one atomic sentence P. Suppose that S is logically omniscient (*i.e.*, S assigns the same credences to logically equivalent statements, and he also assigns zero credence to all contradictions and credence one to all tautologies in his toy language). Thus, all that matters concerning S's coherence (in Joyce's sense) is whether S's credences b in P and $\neg P$ sum to one (and are non-negative). Now, following Joyce, we will associate the truth-value True with the number 1 and the truth-value False with the number 0. Let ϕ be the *numerical* value associated with P's truth-value, and let ψ be the *numerical* value associated with $\neg P$'s truth-value (of course, ϕ and ψ will vary in the obvious ways across the two salient possible worlds: w_1, in which P is false, and w_2, in which P is true). As we saw above, de Finetti's theorem entails that any agent with non-probabilistic credences will be dominated in accuracy (as measured by the Brier score).[111] Rather than describing our "accuracy-dominance reversal theorem" in general terms, we will illustrate it *via* a very simple concrete example, regarding our toy agent S, and assuming the *Brier Score* as our accuracy measure. Suppose that S's credence function (b) assigns the following values P and $\neg P$ (*i.e.*, b entails the following numerical "estimates" of the quantities ϕ and ψ).

	ϕ	ψ
b	$1/2$	$1/4$

TABLE 6.2. The credence function (b) of our simple incoherent agent (S).

de Finetti's theorem entails the existence of a coherent set of estimates (b') of ϕ and ψ, which is more accurate than b (under the Brier Score) in both of the salient possible worlds. We will say that such a b' *Brier-dominates* b *with respect to* ϕ *and* ψ. To make things very concrete, let's look at an example of such a b' in this case. The following table depicts the *Euclidean-closest* such b', relative to Joyce's $\{0, 1\}$-representation of the truth-values (*viz.*, ϕ and ψ). [Ignore the α/β columns of the table, for now — I'll explain their significance, below.]

The estimates entailed by b' are more accurate — with respect to ϕ and ψ — in both w_1 and w_2, according to the Brier Score. A natural question to ask (in light of our discussion of Miller's construction, above) is whether there is a Miller-style symmetric inter-translation that can *reverse* this Brier-dominance relation. Interestingly, it can be shown (proof omitted) that there is *no linear* Miller-style

[111]The various choices of "scoring rule" that one might make in order to render such "accuracy measurements" will not be crucial for the issue that we will be discussing here. The phenomenon will arise for *any* such instantiation of Joyce's framework.

	ϕ	ψ	α	β
b	$1/2$	$1/4$	$9/16$	$3/16$
b'	$5/8$	$3/8$	$3/4$	$1/4$
w_1	0	1	$7/16$	$9/16$
w_2	1	0	$9/16$	$7/16$

TABLE 6.3. An example of the (potential) language-dependence of Joycean Brier-domination

symmetric inter-translation (of the simple form above) that will do the trick. But, there *is* a slightly more complex (non-linear) symmetric inter-translation that will yield the desired reversal (and it is depicted above). Furthermore, it can be shown that *this very same numerical inter-translation* will yield such a reversal for *any* coherent function b' that Brier-dominates b for *this* incoherent agent S (with respect to ϕ and ψ). To be more precise, we have the following theorem about our (particular) agent S.

THEOREM 6.2. *For any coherent function b' that Brier-dominates S's credence function b with respect to ϕ and ψ, there exist quantities α and β that are symmetrically inter-definable with respect to ϕ and ψ,* via *the following* specific *symmetric inter-translations. Where b* Brier-dominates *b' with respect to α and β.*

$$\alpha = \tfrac{1}{2}\phi + \tfrac{1}{2}\psi + \tfrac{1}{16}\left(\tfrac{\phi+\psi}{\phi-\psi}\right) \qquad \beta = \tfrac{1}{2}\phi + \tfrac{1}{2}\psi - \tfrac{1}{16}\left(\tfrac{\phi+\psi}{\phi-\psi}\right)$$

$$\phi = \tfrac{1}{2}\alpha + \tfrac{1}{2}\beta + \tfrac{1}{16}\left(\tfrac{\alpha+\beta}{\alpha-\beta}\right) \qquad \psi = \tfrac{1}{2}\alpha + \tfrac{1}{2}\beta - \tfrac{1}{16}\left(\tfrac{\alpha+\beta}{\alpha-\beta}\right)$$

It is also noteworthy that the *true* values of α and β "behave like truth-values", in the sense that (a) the true value of α (β) in w_1 (w_2) is identical to the true value of β (α) in w_2 (w_1), and (b) the true values of α and β always sum to one. Indeed, these transformations are guaranteed to *preserve coherence* of all dominating b''s, *and* the "truth-vectors".[112]

So, while it is true that there are *some aspects of* "the truth" with respect to which S's credence function b is bound to be less accurate than (various) coherent b''s, it *also seems* to be the case that (for *any* such b') there will be *specifiable, symmetrically inter-definable aspects of* "the truth" on which *the opposite is true* (*i.e.*, with respect to which b' is bound to be less accurate).

Next, we'll consider two possible (general) reactions to this Miller-esque "language dependence of the accuracy of credences" phenomenon. In the end, we think the upshot will be that Joyce needs to tell us more about (precisely) what he

[112]See the companion *Mathematica* notebook for this book (Appendix A), which contains verifications of all of the technical claims made in this section. More general results can be proven (and further constraints can be accommodated on the desired translation scheme). But, all we *need* (dialectically) is *one* incoherent agent S for which we can *ensure* reversals of *all* such Brier-dominance relations *via* a *single*, symmetric inter-translation to/from the ϕ/ψ representation and the α/β representation. See below for further discussion.

means when he says that "credences are (numerical) *estimates* of (numerical) truth-values". Specifically, we think the present phenomenon challenges us to get clearer on the precise content of the *accuracy norm(s)* that are applicable to (or constitutive of) the Joycean cognitive act of "estimation of the (numerical) truth-value of a proposition".

The first response one might have to this phenomenon is that it requires us to maintain that (in some sense) the quantities ϕ and ψ are "more natural" (in this context) than α and β and/or that the "estimation problem" involving ϕ and ψ is somehow "privileged" (in comparison to the α/β "estimation problem"). We don't really see how such an argument would go. First, from the point of view of the α/β-language, the quantities ϕ and ψ seem just as "gerrymandered" as the quantities α and β might appear from the point of view of Joyce's preferred numerical representation of the truth-values. Moreover, there is a disanalogy to the case of physical magnitudes like velocity, since truth-values don't seem to have numerical properties (*per se*). That is, there is already something a little artificial about thinking of truth-values as the sort of things that can be "*numerically estimated*" (where the "estimates" are *numerically scored* for "accuracy").

We think the most promising (and useful) response to the phenomenon is to argue (*i*) that there are crucial *disanalogies* between "estimation" (in Joyce's sense) and "prediction" (in the sense presupposed by Popper and Miller), and (*ii*) these disanalogies imply that our "reversal argument" is presupposing something incorrect regarding the norms appropriate to "estimation".

Here, it is important to note that Joyce does not tell us very much about what he *means* by "estimation". He does say a few things that are suggestive about what "estimation" is *not*. Specifically, Joyce clearly thinks.

(1) Estimates are *not guesses*. Joyce (1998, 587) explicitly distinguishes estimation and guessing. Presumably, guessing (as a cognitive act) doesn't have the appropriate normative structure to ground the sorts of accuracy norms (for credences) that Joyce has in mind.

(2) Estimates are *not expectations*. Joyce (1998, 587–8) explicitly *disavows* thinking of estimates as expectations. Indeed, this is supposed to be one of the novel and distinctive features of Joyce's approach. In fact, it's supposed to be one of the *advantages* of his argument (over previous, similar arguments). Here, Joyce seems to think that expectations have two sorts of (dialectical) shortcomings, in the present context. First, he seems to think that they have a *pragmatic* element, which is not suitable for a non-pragmatic vindication of probabilism. Second, expectations seem to *build-in* a non-trivial amount of *probabilistic structure* [*via* the definition of expectation, which presupposes that $b(\neg p) = 1 - b(p)$], and this makes the assumption that estimates are expectations *question-begging* in the present context.

(3) Estimates are *not assertions that* the values of the parameters *are such-and-so*. This is clear (just from the nature of these "estimation problems"), since it's *not* a good idea to assert things that you know (*a priori*) *must be false*. And, whenever you offer "estimates" of credences that are non-extreme, you know (*a priori*) that the parameters (ϕ and ψ) do *not* take

the values you are offering as estimates (an analogous point can be made with respect to what b and b' "assert" about α and β).

These are the only (definite, precise) commitments about "estimates" that we've been able to extract from Joyce's work (apart from the implicit assumption concerning the appropriateness of "scoring" them in terms of "accuracy" using the Brier score). Unfortunately, these negative claims about what Joyce means by "estimation" do not settle whether our "reversal argument" poses a problem for Joyce.

Let $\ulcorner \mathscr{E}(x,y) = \langle p,q \rangle \urcorner$ be the claim that $\ulcorner S$ is committed to the values $\langle p,q \rangle$ as their "estimates" (in Joyce's sense) of the quantities $\langle x,y \rangle \urcorner$. What we need to know are the conditions under which the following principle (which is implicit in our "reversal argument") is acceptable, given to Joyce's notion of "estimation" ($\mathscr{E}$).

(†) If $\mathscr{E}(\phi,\psi) = \langle p,q \rangle$, then $\mathscr{E}(\alpha,\beta) = f(p,q)$, where f is a symmetric inter-translation function that maps values of $\langle \phi,\psi \rangle$ to/from values of $\langle \alpha,\beta \rangle$.

Presumably, there will be *some* symmetric inter-translation functions f (in *some* contexts) such that (†) *is* acceptable to Joyce. The question is: *which* translation functions f are acceptable — *in our example above*?

Since Joyce doesn't give us a (sufficiently precise) *theory* of $\mathscr{E}$, it is difficult to answer this question definitively. But, if our "reversal argument" is going to be *blocked*, then we presume that Joyce would want to *reject* (†) for *our* inter-translation function $f^\star$ above. It is natural to ask *precisely what grounds* Joyce might have for such a rejection of our $f^\star$.

It is useful to note that (†) is clearly implausible, under *certain* interpretations of $\mathscr{E}$. Presumably, if $\mathscr{E}$ involves *guessing*, then one could argue that (†) should not hold (in general). Perhaps it is just fine for S's guesses about the values of $\langle \phi,\psi \rangle$ to be utterly independent of S's guesses about $\langle \alpha,\beta \rangle$ (at least, to the extent that we understand the "norms of guessing"). Similarly, if $\mathscr{E}$ involves *expectation*, then (†) will *demonstrably fail* (in general) for *non*-linear functions like our $f^\star$. [Although, on an *expectation* reading of "estimate," (†) will *demonstrably hold* for all *linear* inter-translations.[113]] Unfortunately for Joyce, neither of these interpretations of $\mathscr{E}$ is available to him. So, this yields no concrete reasons to reject (†) in our example.

On the other hand, if $\mathscr{E}$ involved *assertion* (as in item 3 above), then (†) would seem to be more plausible. On an assertion reading of $\mathscr{E}$, (†) is tantamount to a simple form of deductive closure for assertoric commitments (in the traditional

[113] Pedersen and Glymour (2012) provide a response to the worries expressed in this section. They point out that the properties of expectation will be *guaranteed* to hold, as long as a proper scoring rule is used in the Joycean argument. In this sense, their response to the present worries is to insist on the propriety of measures of credal inaccuracy. Doing so ensures that $\mathscr{E}$ yields (only) *probabilistically coherent* commitments. One of the reasons we were (initially) worried about this language dependence argument is that we took it to be a (potential) reason to worry about *propriety itself*. However, our present arguments — which we think also give some reason to *adopt propriety* — do not fall prey to these language dependence worries. Indeed, this is one of the reasons we have since adopted the approach described in this book. Moreover, unless a "purely alethic" argument for *propriety* can be given, we remain unsure whether "veritists" (*i.e.*, "truth monists") like Joyce (2013) and Pettigrew (2013a) will be able to avoid this (possible) threat of language dependence. We will discuss the prospects of this sort of appeal to "accuracy-first epistemology" below.

sense).[114] And, this would be very similar to the way Popper and Miller were thinking about the *predictions* of (deterministic, quantitative) scientific *theories*. It seems clear that $\mathscr{E}$ is *not exactly like that* (in this context), but this (alone) doesn't give us any concrete reasons to *reject* (†) in this case.

We submit that what is needed here is a (sufficiently precise) *theory* of $\mathscr{E}$, which satisfies Joyce's explicit commitments (1)–(3) above, and which is also precise enough to *explain why* (†) *should fail for* $f^\star$ (in our example above).

3.3. Caie's (Original) Example. In Part I (Negative Phase), we discussed the full belief analogue of Caie's (2013) proposed counterexample to Joyce's argument for probabilism. In this section, we'll discuss Caie's original example. Consider the following self-referential claim regarding an agent S's credence function b.

(P) $b(P) < 1/2$.

In words, P says (of itself) that S assigns a credence less than $1/2$ to P. Consider the agenda $\mathcal{A} = \{P, \neg P\}$. As usual, we assume our agent is logically omniscient, *etc.* So, S's credence function b is probabilistic iff b satisfies the additivity constraint $b(P) + b(\neg P) = 1$. If we adopt the Brier score as our measure of credal inaccuracy, then we can argue that *the only non-Brier-dominated* credal assignment on $\mathcal{A}$ is the *non-probabilistic* assignment $b(P) = 1/2$; $b(\neg P) = 1$.

The argument for this claim is analogous to the argument (Part I, Negative Phase) that the only non-dominated belief set in the full belief analogue of Caie's case is the probabilistically unrepresentable set $\{B(P), B(\neg P)\}$. Here's one way to make the argument, *via* minimization of Brier score. Suppose that S's credal assignment is $b(P) = \alpha$ and $b(\neg P) = \beta$. Then, or so Caie's argument goes, *the only possible worlds* that are relevant for evaluating b are the ones that are not ruled-out semantically by the definition of P (together with the attitude taken by S). Table 6.4 below is similar to Table 2.2 from Part I (p. 39). The "×"s indicate that these "possible worlds" are ruled-out (semantically) by the definition of P (together with the attitude taken by S). In the worlds w that are not semantically ruled out, the cells of the table contain the salient Brier score $\mathcal{I}_b(b, w)$.

	P	$\neg P$	$b(P) < 1/2$	$b(P) \geq 1/2$
w_1	**F**	**T**	×	$\mathcal{I}_b(b, w_1) = \alpha^2 + (\beta - 1)^2$
w_2	**T**	**F**	$\mathcal{I}_b(b, w_2) = (\alpha - 1)^2 + \beta^2$	×

TABLE 6.4. Caie's original example

Thus, the (total) Brier score of b (at the single, non-ruled out possible world) can be expressed in terms of α and β *via* the following piecewise function.

$$\text{Brier score of } b = \begin{cases} \alpha^2 + (\beta - 1)^2 & \text{if } P \text{ is true (\textit{i.e.}, if } b(P) < 1/2\text{)}, \\ (\alpha - 1)^2 + \beta^2 & \text{if } P \text{ is false (\textit{i.e.}, if } b(P) \geq 1/2\text{)}. \end{cases}$$

[114]Even on this reading, (†) is not very plausible *to us*, given the views about full belief (and assertion) that we articulated and Part I.

Assuming $\alpha, \beta \in [0, 1]$, this piecewise function has a *unique minimum* at $\langle a, b \rangle = \langle 1/2, 1 \rangle$. It follows that $b(P) = 1/2$; $b(\neg P) = 1$ is the unique credal assignment that is non-dominated in Brier score. And, as in the full belief version of the example, this assignment *violates our* (probabilistic) *evidential requirement* ($\mathcal{R}_b$).

Everything we said in Part I about the full belief version of the example carries over (*mutatis mutandis*) to Caie's original credal case. So, we won't repeat the bulk of that dialectic (including the ramifications for "epistemic decision theory/epistemic teleology") here. However, we would like to elaborate upon a worry we voiced in passing (*fn.* 64) regarding the evidential propriety of the measures of inaccuracy used to analyze Caie-style cases. In our analyses of each of the Caie-style cases (Parts I and III), we used measures of inaccuracy that were *evidentially proper* in *ordinary* (*viz.*, *non*-Caie) cases. And, in both cases, these same measures turned out to be *evidentially improper*, since being probabilistically representable no longer ensured being non-dominated according to these inaccuracy measures. In the credal case, the dialectic is especially subtle, since (presumably) Caie is offering his examples as *reasons to reject probabilism* ($\mathcal{R}_b$) as an evidential requirement for credences. We wonder whether that is the right reaction to the example.

We conjecture that *no* inaccuracy measure (at least, none that could reasonably be taken to be a *divergence*) can be *proper*, when applied to Caie's example.[115] One way to react to this sort of impossibility is to *reject probabilism* ($\mathcal{R}_b$) as an evidential requirement (and that would undermine the approach taken in this book). But, another reaction is to say that examples such as these are *outside the proper domain of application* of the present framework, since the framework only applies in contexts where an evidentially proper inaccuracy measure can be defined. In the case of comparative confidence, we pointed out that *no evidentially proper inaccuracy measure* can be defined for the stronger notion of "accuracy" implied by the assumption of a *unique* "vindicated" comparative confidence ordering $\mathring{\succeq}_w$. Our reaction to *that* impossibility was *not* to reject our probabilistic evidential requirement ($\mathcal{R}_{\succeq}$). Rather, it was to formulate a *weaker* notion of "accuracy" (and a *stronger* corresponding notion of "inaccuracy") and then show that we *could* successfully run our explication of "coherence" using that weaker (stronger) notion. The prospects for a similar move here seem rather dim, as it seems that *no* proper scoring rule is definable in this case (even if we somehow revise our notion of "accuracy"). However, this may just mean that a more radical revision of the framework is necessary.[116] Even if it is not possible to take an analogous tack here, it seems that we still have the option of viewing these sorts of examples as outside the proper domain of application of our framework. Perhaps cases like

---

[115] That is, it seems that (holding fixed the classical underlying formal semantical structure — *fn.* 116) *no function of* α *and* β *that is a divergence* between $\langle \alpha, \beta \rangle$ and the world vector of the (suitable) non-ruled out world w can be proper in the context of Caie's example. We don't have a proof of this, but we state is as a conjecture, since we suspect any such function will either (a) fail to satisfy all the axioms for divergences (Csiszár, 1991) or (b) fail to be proper.

[116] Perhaps a non-classical underlying semantics will be required to properly model these sorts of cases. This wouldn't be too surprising, since allowing self-reference into the language we use to describe credal states can lead to paradox in the presence of other assumptions (*e.g.*, introspection assumptions). Maybe a non-classical semantics which avoids such paradoxes would allow us to create a more appropriate formal framework for modeling such cases (Campbell-Moore, 2014a). That is, maybe there is a way of constructing a proper scoring rule within a suitable non-classical generalization of our framework. This is a crucial open theoretical question regarding Caie's example.

these just aren't in the spirit of our evidentialist slogan: "(numerical) probabilities reflect evidence." That is, perhaps it is unclear what the evidential relations *are* in such cases (and not simply whether they can be modeled probabilistically).[117]

Indeed, there seems to be a pathological kind of *evidential instability* involved in Caie-style cases. To see this, ask yourself what the evidential norm for credences recommends for S's credences regarding P in the following two cases.

Case 1. S's total evidence supports P.

Case 2. S's total evidence does not support P.

In Case 1, it is natural to suppose that S should (according to the evidential norm for credences) assign a credence greater than $1/2$ to P. But, doing so would make P *false* (and S knows this), which would seem to contradict the assumption that S's total evidence supports P. Similarly, in Case 2, it is natural to suppose that S should (according to the evidential norm for credences) assign a credence less than $1/2$ to P. But, doing so would make P *true* (and S knows this), which would seem to contradict the assumption that S's total evidence does not support P. It's not obvious that the correct response to this sort of evidential instability is to reject (WADA_b)/probabilism as a coherence requirement for credences.[118]

One final issue is worth discussing in connection with Caie-style cases. Let's abstract away from numerical credences here and try to formulate a *comparative confidence* rendition of Caie's example. The most obvious candidate would be:

(Q) $\neg Q \succ Q$.

Interestingly, (Q) does *not* seem to pose analogous problems for our explication of "coherence" for comparative confidence relations (at least, we haven't been able to derive anything using Q that causes Caie-style trouble our approach). It is an open question whether a problematic Caie-style example can be constructed for comparative confidence.

4. Epistemic Value Monism, Pluralism, Propriety and Coherence

The general approach to grounding (formal, synchronic, epistemic) coherence requirements taken in this monograph can be summarized in the following way.

(1) We *assume* that "(numerical) probabilities reflect evidence". That is, we assume that, in each epistemic context, there is *some* (evidential) probability function which reflects the evidential relations in that context. This implies a common *evidential* requirement for judgments of various types: *representability by some probability function*, where the appropriate sense

[117] It is interesting to note that the Dutch Book argument does not seem to lose its force in Caie cases (Campbell-Moore, 2014b). This is an independent (albeit pragmatic) reason to suspect that Caie's argument does not actually refute probabilism as a coherence requirement for credences.

[118] One can write down self-referential statements that engender an internal conflict between the Truth Norm and the Evidential norm for belief. Consider the following claim:

(X) S's total evidence does not support X.

Should S believe X? Well, S's total evidence supports X *iff* X is false. So, it's unclear what S's attitude toward X should be. But, is this a compelling reason to reject the Evidential Norm for belief (*i.e.*, the norm that says one should believe a claim iff one's total evidence supports that claim)?

of "representability" depends on the judgment type. Specifically, this lead to the three $\mathcal{R}$-requirements discussed in the book: ($\mathcal{R}$), ($\mathcal{R}_{\succeq}$) and ($\mathcal{R}_b$).

(2) We *assume* an *extensional* notion of "inaccuracy" for judgments, which supervenes on the truth-values of their contents. This leads to extensional alethic norms and requirements, which are analogous to the Truth Norm and the Consistency Requirement for belief. Specifically, we have notions of actual and possible vindication for each of the three types of judgments discussed in the book: (AV), (PV), ($\text{AV}_{\succeq}$), ($\text{PV}_{\succeq}$) and (AV_b), (PV_b).

(3) We define an (additive) *evidentially proper inaccuracy measure*, relative to the assumptions in (1) and (2). Specifically, we have seen (additive) evidentially proper inaccuracy measures for each of our three judgment types: $\mathcal{I}$, $\mathcal{I}_{\succeq}$ and $\mathcal{I}_b$.

(4) Steps (1)–(3) give rise to *accuracy dominance avoidance* requirements, which are entailed by both the alethic and the evidential norms (*i.e.*, these are *conflict-proof* requirements in the sense of our desideratum $\mathcal{D}$ from Part I) for the judgment type in question. Specifically, we have derived both weak and strict accuracy dominance avoidance requirements for each of our three judgment types: (WADA), (SADA), ($\text{WADA}_{\succeq}$), ($\text{SADA}_{\succeq}$) and (WADA_b), (SADA_b).

We think this approach yields an illuminating explication of "coherence" for each of the three judgment types discussed in the book. In each case, the accuracy-dominance avoidance requirements are (plausibly) necessary requirements of ideal epistemic rationality. In the case of full belief, we offered the probabilistic representability requirement ($\mathcal{R}$) *itself* as a *stronger* coherence requirement (Figure1.2). The comparative confidence case has a similar structure (Figure 3.4), with probabilistic representability ($\mathcal{R}_{\succeq}$) being *strictly stronger* than the requirement of accuracy dominance avoidance ($\text{WADA}_{\succeq}$). In the credal case, the proposal of probabilistic representability as a coherence requirement may seem "circular", since in the credal case probabilistic representability ($\mathcal{R}_b$) happens to be *equivalent* to the accuracy dominance avoidance requirement (WADA_b). We do not think this renders our explication of credal coherence "circular" (or otiose). The credal case just happens to be the *limiting case* of a general theoretical explication of coherence. This is why we wrote the book "backward" (in a historical sense).

One needs to see the applications of the framework in the right order — *from most coarse-grained to most fine-grained* — in order to appreciate what is really going on with the "arguments for probabilism" that inspired the framework in the first place. We think the best way to view these explications of credal coherence is not as "arguments for probabilism" (*per se*), but rather as *limiting cases* of the more general explication of "coherence" (in terms of accuracy dominance avoidance, relative to an evidentially proper inaccuracy measure) that we have offered in this book. This perspective is not shared by all authors currently writing on these topics. The prevailing view now seems to be that arguments like Joyce's are best viewed as *arguments for probabilism*, from an "accuracy-first" perspective.

The "accuracy-first" or "veritist" perspective is a popular one in contemporary epistemology (Goldman, 1999). More generally, *epistemic value monism* has received a lot of attention in the contemporary literature (Zagzebski, 1994; Sosa, 2003). The idea behind "accuracy-first" epistemology (which I will hereafter refer to as *veritism*) is that *all epistemic value supervenes on the alethic* (*i.e.*, on the *truth-values* of the contents of the judgments being epistemically evaluated). From this veritist perspective, *justification* (or evidential support) is (at best) *merely instrumentally* epistemically valuable. If believing things that are supported by your evidence helps promote the aim of true belief, then it is — to that extent — *instrumentally* valuable. But, believing in accordance with your evidence is not epistemically valuable *per se*. On the other extreme, we have *evidentialists*, who seem to think that believing in accordance with your evidence is the *sine qua non* of epistemic achievement (Clifford, 1877; Conee and Feldman, 2004; Kolodny, 2008). In this sense, the debate between veritists and evidentialists would seem to be an *intramural* dispute among epistemic value monists. We favor a *pluralist* stance, which takes *both* alethic *and* evidential considerations to be crucial when it comes to the epistemic evaluation of doxastic states (DePaul, 2001). Indeed, we think that hallmark of epistemic coherence requirements is that they are *points of convergence* (or *points of agreement*) between alethic and evidential norms.

The veritist has a different way of thinking about coherence requirements (Joyce, 2013; Pettigrew, 2013a). According to the veritist, we must ground coherence requirements (*e.g.*, probabilism) in a *purely alethic* way — without any (ineliminable) appeal to evidential support, or justification (or any other putative extra-alethic epistemic considerations). In order to do this, the veritist must try to provide some "purely alethic" way of grounding the use of *proper* measures of doxastic inaccuracy (as opposed to improper ones like Maher's (2002) inaccuracy measure L_1). As we explained above, Joyce (2009) explicitly appeals to *evidential* considerations to defend his use of proper inaccuracy measures. Of course, we are sympathetic to this approach. Indeed, this is why we have adopted the term *evidentially proper* to refer to this crucial property of inaccuracy measures. As far as we are concerned, propriety is important *precisely because it ensures the alignment of evidential and alethic norms and requirements*. We remain skeptical about the prospects of providing a "purely alethic" argument for the use of proper inaccuracy measures. Indeed, we think the attempt to provide such "purely alethic" ways of grounding propriety (or coherence) is misguided, since coherence requirements are best understood as formal, wide-scope requirements that are entailed by both alethic and evidential considerations. Be that as it may, it seems to us that the challenge of providing a "purely alethic" argument for propriety (or, more generally, for coherence requirements such as probabilism) has not yet been met.[119]

5. What if there are no coherence requirements?

We close with a brief discussion of a challenge to the *very existence* of (formal, epistemic) coherence requirements that has been articulated by Kolodny (2007, 2008). According to Kolodny, *the only norms* there are in epistemology are (narrow-scope) *evidentialist* norms, which require us form individual judgments that are supported by the total evidence. On this view, *there are no (wide-scope) coherence requirements*. There are wide-scope *constraints* on judgment sets that *emerge* from the evidential norms, but these constraints are mere "epistemic epiphenomena." In other words, the *only* thing wrong with someone who is "incoherent" (in

[119]Pettigrew (2014) sketches an argument for propriety that is meant to meet this challenge. We have not yet had a chance to analyze this argument in detail. We aim to do so in a future draft.

our sense) is that they are (actually) violating the (narrow-scope) evidential norm. That is to say, being "incoherent" (in our sense) is not an epistemic defect *per se*. While Kolodny's attack on coherence requirements is launched from a (monist) *evidentialist* perspective, one can imagine a similar coherence-eliminativist argument, which starts from a (monist) *veritist* point of view. A veritist could also maintain that there are no (wide-scope) coherence requirements — there are only (narrow-scope) *alethic norms*. And, what *appear* to be coherence requirements are mere *constraints* on judgment sets that *emerge* from the alethic norms (*viz.*, the only epistemic mistakes are actual violations of the narrow-scope alethic norms).

In contrast, the stance we have taken in this book is a pluralist one, which emphasizes the importance of both alethic and evidential norms. From this point of view, the problem with an incoherent agent (in our sense) is that they (ideally) are in a position to know (*a priori*) that they are in violation of both the evidential and the alethic norms. We think this is a hallmark of epistemic irrationality. That is to say, we think coherence (in our sense) is helpful in understanding the interplay between the alethic and the evidential in our evaluations of the epistemic rationality (or irrationality) of doxastic states.

APPENDIX A

Companion *Mathematica* Notebook

> . . . the trouble about arguments is, they ain't nothing but theories, after all, and theories don't prove nothing.
>
> Mark Twain

This Appendix contains proofs of many of the theorems in the book. However, some of the theorems (especially those in Part II) were verified using *Mathematica*. Rather than reproduce the relevant bits of *Mathematica* code here, I have created the following companion *Mathematica* notebook to go along with the book.[120]

`http://fitelson.org/coherence.nb/`

We have also created a PDF version of the companion notebook, for those who do not have access to *Mathematica*.

`http://fitelson.org/coherence.nb.pdf`.

The companion notebook contains three sections — one for each of the three Appendices of the book. And, the verifications presented in the notebook are presented in the same order that they appear in the Appendices.

[120]Not only does the *Mathematica* notebook allow one to verify all of the corresponding technical claims in the Appendices, but the (rather vast) suite of *Mathematica* functions developed in the notebook can be used to further explore the various models presented in this book.

APPENDIX B

Theorems from Part I

Proof of Theorem 1.1.
[**B** is non-dominated iff (⇔) **B** contains no witnessing set.]

(⇒) We prove the contrapositive. Suppose that $\mathbf{S} \subseteq \mathbf{B}$ is a witnessing set. Let $\mathbf{B}'$ agree with **B** on all judgments outside **S** and disagree with **B** on all judgments in **S**. By the definition of a witnessing set, $\mathbf{B}'$ must weakly dominate **B** in accuracy [$\mathcal{I}(\mathbf{B}, w)$]. Thus, **B** is dominated.

(⇐) We prove the contrapositive. Suppose that **B** is dominated, *i.e.*, that there is some $\mathbf{B}'$ that weakly dominates **B** in accuracy [$\mathcal{I}(\mathbf{B}, w)$]. Let $\mathbf{S} \subseteq \mathbf{B}$ be the set of judgments on which **B** and $\mathbf{B}'$ disagree. Then, **S** is a witnessing set.

Proof of Theorem 1.2.
[**B** is non-dominated if (⇐) there is a probability function Pr that represents **B**.]

Let Pr be a probability function that represents **B** in sense of Definition 1.2. Consider the expected inaccuracy, as calculated by Pr, of a belief set **B** — the sum over all worlds w of $\Pr(w) \cdot \mathcal{I}(\mathbf{B}, w)$. Since $\mathcal{I}(\mathbf{B}, w)$ is a sum of components for each proposition (1 if **B** disagrees with w on the proposition and 0 if they agree), and since expectations are linear, the expected inaccuracy is the sum of the expectation of these components. The expectation of the component for disbelieving p is $\Pr(p)$ while the expectation of the component for believing p is $1 - \Pr(p)$. Thus, if $\Pr(p) > 1/2$ then believing p is the attitude that uniquely minimizes the expectation, while if $\Pr(p) < 1/2$ then disbelieving p is the attitude that uniquely minimizes the expectation. Thus, since Pr represents **B**, this means that **B** has strictly lower expected inaccuracy than any other belief set with respect to Pr (*i.e.*, that **B** uniquely minimizes expected inaccuracy, relative to Pr). Suppose, for *reductio*, that some $\mathbf{B}'$ (weakly) dominates **B**. Then, $\mathbf{B}'$ must be no more inaccurate than **B** in any world, and thus $\mathbf{B}'$ must have expected inaccuracy no greater than that of **B**. But **B** has strictly lower expected inaccuracy than any other belief set. Contradiction. Therefore, no $\mathbf{B}'$ can dominate **B**, and so **B** must be non-dominated.

Verifications of the Claims that (SADA) $\not\Rightarrow$ (WADA) and (NDB) $\not\Rightarrow$ ($\mathcal{R}$).

Consider a sentential language $\mathcal{L}$ with two atomic sentences X and Y. The Boolean algebra $\mathcal{B}$ generated by $\mathcal{L}$ contains 16 propositions (corresponding to the subsets of the set of four state descriptions of $\mathcal{L}$, *i.e.*, the set of four salient possible worlds). Table C.2 depicts $\mathcal{B}$, and two opinionated belief sets ($\mathbf{B}_1$ and $\mathbf{B}_2$) on $\mathcal{B}$.
We have the following four salient facts regarding $\mathbf{B}_1$ and $\mathbf{B}_2$:

$\mathcal{B}$	$\mathbf{B}_1$	$\mathbf{B}_2$
$\neg X \,\&\, \neg Y$	D	D
$X \,\&\, \neg Y$	D	D
$X \,\&\, Y$	D	D
$\neg X \,\&\, Y$	D	D
$\neg Y$	D	D
$X \equiv Y$	D	D
$\neg X$	B	B
X	B	D
$\neg(X \equiv Y)$	D	D
Y	D	D
$X \vee \neg Y$	B	B
$\neg X \vee \neg Y$	B	B
$\neg X \vee Y$	B	B
$X \vee Y$	D	B
$X \vee \neg X$	B	B
$X \,\&\, \neg X$	D	D

TABLE B.1. Examples showing (SADA) $\not\Rightarrow$ (WADA) and (NDB) $\not\Rightarrow$ ($\mathcal{R}$)

(1) $\mathbf{B}_1$ is weakly dominated (in accuracy) by belief set $\mathbf{B}_2$. Thus, $\mathbf{B}_1$ violates (NDB)/(WADA).

(2) $\mathbf{B}_1$ is not strictly dominated (in accuracy) by any belief set over $\mathcal{B}$. Thus, $\mathbf{B}_1$ satisfies (SADA).

(3) $\mathbf{B}_2$ is not weakly dominated (in accuracy) by any belief set over $\mathcal{B}$. Thus, $\mathbf{B}_2$ satisfies (WADA).

(4) $\mathbf{B}_2$ is not represented (in the sense of Definition 1.2) by any probability function on $\mathcal{B}$. Thus, $\mathbf{B}_2$ violates ($\mathcal{R}$).

Claims (1)–(4) are all verified in the companion *Mathematica* notebook.

PROOF OF THEOREM 1.3.
[For all $n \geq 2$ and for each set of propositions $\mathbf{P} \in \mathbb{B}_n$, if $r \geq \frac{n-1}{n}$ then ($\mathcal{R}_r$) rules out believing every member of $\mathbf{P}$, while if $r < \frac{n-1}{n}$, then ($\mathcal{R}_r$) doesn't rule out believing every member of $\mathbf{P}$.]

Let $\mathbf{P}$ be a member of $\mathbb{B}_n$, *i.e.*, $\mathbf{P}$ consists of n propositions, there is no world in which all of these n propositions are true, but for each proper subset $\mathbf{P}' \subset \mathbf{P}$ there is a world in which all members of $\mathbf{P}'$ are true.

Let $\phi_1, \ldots, \phi_n$ be the n propositions in $\mathbf{P}$. Let each w_i be a world in which ϕ_i is false, but all other members of $\mathbf{P}$ are true. Let Pr be the probability distribution that assigns probability $1/n$ to each world w_i and 0 to all other worlds. If $r < \frac{n-1}{n}$, then

Pr shows that the belief set $\mathbf{B_P} := \{B(\phi_1), \ldots, B(\phi_n)\}$, which includes the belief that ϕ_i for each $\phi_i \in \mathbf{P}$, satisfies $(\mathcal{R}_r)$. This establishes the second half of the theorem.

For the first half, we will proceed by contradiction. Thus, assume that $\mathbf{P}$ is a member of $\mathbb{B}_n$ such that the belief set $\mathbf{B_P} := \{B(\phi_1), \ldots, B(\phi_n)\}$, which includes the belief that ϕ_i for each $\phi_i \in \mathbf{P}$, is *not* ruled out by $(\mathcal{R}_{n-1/n})$. Then there must be some Pr such that for each i, $\Pr(\phi_i) > {}^{n-1}/{}_n$. This means that for each i, $\Pr(\neg\phi_i) < {}^1/{}_n$. Since the disjunction of finitely many propositions is at most as probable as the sum of their individual probabilities, this means that $\Pr(\neg\phi_1 \vee \ldots \vee \neg\phi_n) < 1$. But since $\mathbf{P}$ is inconsistent, $\neg\phi_1 \vee \ldots \vee \neg\phi_n$ is a tautology, and therefore must have probability 1. This is a contradiction, so $\mathbf{B_P}$ must be ruled out by $(\mathcal{R}_{n-1/n})$.

PROOF OF THEOREM 1.4.
[**B** is not strictly dominated iff ($\Leftrightarrow$) **B** contains no witnessing$_1$ set.
That is: (SADA) $\Leftrightarrow$ (NW$_1$S)]

($\Rightarrow$) We'll prove the contrapositive. Suppose that $\mathbf{S} \subseteq \mathbf{B}$ is a witnessing$_1$ set. Let $\mathbf{B}'$ agree with $\mathbf{B}$ on all judgments outside $\mathbf{S}$ and disagree with $\mathbf{B}$ on all judgments in $\mathbf{S}$. By the definition of a witnessing$_1$ set, $\mathbf{B}'$ must strictly dominate $\mathbf{B}$ in accuracy $[\mathcal{I}(\mathbf{B}, w)]$. Thus, $\mathbf{B}$ is strictly dominated.

($\Leftarrow$) We prove the contrapositive. Suppose $\mathbf{B}$ is strictly dominated, *i.e.*, that there is some $\mathbf{B}'$ that strictly dominates $\mathbf{B}$ in accuracy $[\mathcal{I}(\mathbf{B}, w)]$. Let $\mathbf{S} \subseteq \mathbf{B}$ be the set of judgments on which $\mathbf{B}$ and $\mathbf{B}'$ disagree. Then, $\mathbf{S}$ is a witnessing$_1$ set.

PROOF OF THEOREM 1.5.
[**B** is probabilistically representable (in the sense of Definition 1.2) only if **B** contains no witnessing$_2$ set. That is, $(\mathcal{R}) \Longrightarrow$ (NW$_2$S).]

In our proof of Theorem 1.2, we established that if Pr represents **B**, then **B** has strictly lower expected inaccuracy than any other belief set with respect to Pr. Assume, for *reductio*, that $\mathbf{S} \subseteq \mathbf{B}$ is a witnessing$_2$ set for **B**. Let $\mathbf{B}'$ agree with **B** on all judgments outside **S** and disagree with **B** on all judgments in **S**. Then by the definition of a witnessing$_2$ set, $\mathbf{B}'$ must be no more inaccurate than **B** in any world. But this contradicts the fact that **B** has strictly lower expected inaccuracy than $\mathbf{B}'$ with respect to Pr. So the witnessing$_2$ set must not exist.

PROOF OF THEOREM 1.6.
[**B** is probabilistically representable (in the sense of Definition 1.2) only if **B** satisfies both (NDB) and (NCP). That is, $(\mathcal{R}) \Longrightarrow$ (NDB & NCP).]

Theorem 1.2 implies $(\mathcal{R}) \Longrightarrow$ (NDB). And, it is obvious that $(\mathcal{R}) \Longrightarrow$ (NCP), since no probability function can probabilify both members of a contradictory pair and no probability function can dis-probabilify both members of a contradictory pair.

COUNTEREXAMPLE TO THE CONVERSE OF THEOREM 1.6.
[$(\mathcal{R}) \not\Leftarrow$ (NDB & NCP).]

Let there be six possible worlds, $w_1, w_2, w_3, w_4, w_5, w_6$. Consider the agenda $\mathcal{A}$ consisting of the following four propositions (*i.e.*, $\mathcal{A} \overset{\text{def}}{=} \{p_1, p_2, p_3, p_4\}$).

$$p_1 = \{w_1, w_2, w_3\}$$

$$
\begin{aligned}
p_2 &= \{w_1, w_4, w_5\} \\
p_3 &= \{w_2, w_4, w_6\} \\
p_4 &= \{w_3, w_5, w_6\}
\end{aligned}
$$

Let $\mathbf{B} \overset{\text{def}}{=} \{B(p_1), B(p_2), B(p_3), B(p_4)\}$. $\mathbf{B}$ is itself a witnessing$_2$ set, since, in every possible world, exactly two beliefs (*i.e.*, exactly half of the beliefs) in $\mathbf{B}$ are accurate. So by Theorem 5, $\mathbf{B}$ is not probabilistically representable. However, $\mathbf{B}$ satisfies (NDB). To see this, note that every belief set on $\mathcal{A}$ has an expected inaccuracy of 2, relative to the uniform probability distribution. This implies that no belief set on $\mathcal{A}$ dominates any other belief set on $\mathcal{A}$. Finally, $\mathbf{B}$ satisfies (NCP), since every pair of beliefs in $\mathbf{B}$ is consistent.[121]

PROOF OF THEOREM 1.7.
[(SSADA) ⇔ (NCB).]

(⇒) Suppose $\mathbf{B}$ violates (NCB). That is, suppose $\mathbf{B}$ contains $B(\bot)$. Consider the alternative belief set $\mathbf{B}'$ which agrees with $\mathbf{B}$ everywhere *except* regarding $\bot$. In every possible world w, $\mathbf{B}'$, $\mathbf{M}(\mathbf{B}', \mathring{\mathbf{B}}_w) = \mathbf{M}(\mathbf{B}, \mathring{\mathbf{B}}_w) \setminus B(\bot) \subset \mathbf{M}(\mathbf{B}, \mathring{\mathbf{B}}_w)$. So, $\mathbf{B}$ violates (SSADA). A parallel argument establishes the $B(\top)$ case.

(⇐) Suppose $\mathbf{B}$ violates (SSADA). And, suppose, for *reductio*, that $\mathbf{B}$ *satisfies* (NCB). That is, $\mathbf{B}$ contains no inaccurate judgments regarding logically non-contingent claims. Because $\mathbf{B}$ violates (SSADA), there exists a $\mathbf{B}'$ such that $\mathbf{M}(\mathbf{B}', \mathring{\mathbf{B}}_w) \subset \mathbf{M}(\mathbf{B}, \mathring{\mathbf{B}}_w)$, in every possible world w. Then, $\mathbf{B}'$ and $\mathbf{B}$ must agree on all logically non-contingent claims. And, there must be some contingent claim p on which $\mathbf{B}'$ and $\mathbf{B}$ disagree. For any such claim p, there will be some worlds w_p in which p is true and some $w_{\neg p}$ in which p is false. Suppose $\mathbf{B}$ contains $B(p)$. Then, $\mathbf{B}'$ contains $D(p)$. In this case, $\mathbf{M}(\mathbf{B}', \mathring{\mathbf{B}}_{w_p}) \not\subset \mathbf{M}(\mathbf{B}, \mathring{\mathbf{B}}_{w_p})$, since $D(p)$ is a mistaken judgment (in w_p) that is contained in $\mathbf{B}'$, but not in $\mathbf{B}$. On the other hand, if $\mathbf{B}$ contains $D(p)$, then $\mathbf{B}'$ must contain $B(p)$, and $\mathbf{M}(\mathbf{B}', \mathring{\mathbf{B}}_{w_{\neg p}}) \not\subset \mathbf{M}(\mathbf{B}, \mathring{\mathbf{B}}_{w_{\neg p}})$, since $B(p)$ is a mistaken judgment (in $w_{\neg p}$) that is contained in $\mathbf{B}'$, but not in $\mathbf{B}$. Contradiction. So, $\mathbf{B}$ must *violate* (NCB).

[121] We can extend this counterexample to an example that isn't restricted to $\mathcal{A}$, but in fact is opinionated across the whole algebra $\mathcal{B}$ of propositions constructible out of the six possible worlds. Here's how to extend $\mathbf{B}$ to a belief set $\mathbf{B}'$ over $\mathcal{B}$. For the propositions p that are true in four or more worlds, $\mathbf{B}'$ contains $B(p)$. For the propositions q that are true in two or fewer worlds, $\mathbf{B}'$ contains $D(q)$. For each pair of complementary propositions that are true in exactly three worlds, $\mathbf{B}'$ includes belief in one and disbelief in the other (*i.e.*, $\mathbf{B}'$ may include any consistent pair of attitudes toward each pair of complementary propositions that are true in exactly three worlds). By construction, this set $\mathbf{B}'$ satisfies (NCP). Additionally, the belief set $\mathbf{B}'$ is tied for minimal expected inaccuracy on $\mathcal{B}$, relative to the uniform probability distribution on $\mathcal{B}$. Therefore, $\mathbf{B}'$ satisfies (NDB).

APPENDIX C

Theorems from Part II

PROOF OF THEOREM 3.1
[(WADA$_{\succeq}$) $\Rightarrow$ ($\mathfrak{C}$)]

Suppose $\succeq$ violates ($\mathfrak{C}$). There are only two ways this can happen. Either $\succeq$ violates (A_1) or $\succeq$ violates (A_2).

Suppose $\succeq$ violates (A_1). Because $\succeq$ is total, this means $\succeq$ is such that $\bot \succeq \top$. Consider the relation $\succeq'$ which agrees with $\succeq$ on all comparisons outside the $\langle \bot, \top \rangle$-fragment, but which is such that $\top \succ' \bot$. Because $\top \,\&\, \neg\bot$ *is true in every possible world* w, it follows that $\mathbf{M}(\succeq', w) \subset \mathbf{M}(\succeq, w)$ *in every possible world* w. Thus, (A_1) actually follows from (SSADA$_{\succeq}$). And, of course, (WADA$_{\succeq}$) entails (SSADA$_{\succeq}$).

Suppose $\succeq$ violates (A_2). Because $\succeq$ is total, this means there is a pair of propositions p and q in $\mathcal{A}$ such that (a) p entails q but (b) $p \succ q$. Consider the relation $\succeq'$ which agrees with $\succeq$ on every judgment *except* (b), and which is such that $q \succ' p$. Table C.1 depicts the $\langle p, q \rangle$-fragment of the relations $\succeq$ and $\succeq'$ in the three salient possible worlds (the second row/world is *impossible*, since p entails q). By (b) and (LO), p and q are not logically equivalent. Thus, world w_2 is a live possibility. Therefore, $\succeq'$ weakly $\mathcal{I}_{\succeq}$-dominates $\succeq$.

w_i	p	q	$\succeq$	$\succeq'$	$\mathcal{I}_{\succeq}(\succeq, w_i)$	$\mathcal{I}_{\succeq}(\succeq', w_i)$
w_1	T	T	$p \succ q$	$q \succ' p$	0	0
	T	F				
w_2	F	T	$p \succ q$	$q \succ' p$	2	0
w_3	F	F	$p \succ q$	$q \succ' p$	0	0

TABLE C.1. The $\langle p, q \rangle$-fragments of $\succeq$ and $\succeq'$ in the 3 salient possible worlds [(A_2) case of Theorem 3.1].

VERIFYING THEOREM 3.3
[(WADA$_{\succeq}$) $\Rightarrow$ (A_3)]

Suppose $\succeq$ violates (A_3). Because $\succeq$ is total, this means there must exist $p, q, r \in \mathcal{A}$ such that (a) $p \vDash q$, (b) $\langle q, r \rangle$ are mutually exclusive, (c) $q \succ p$, but (d) $p \vee r \succeq q \vee r$. Let $\succeq'$ agree with $\succeq$ on every judgment, *except* (d). That is, let $\succeq'$ be such that (e) $q \succ' p$ and (f) $q \vee r \succ' p \vee r$. There are only four worlds compatible with the precondition of (A_3), and these are depicted in Table C.2, below.

p	q	r	possible world w_i
T	**T**	**T**	
T	**T**	**F**	w_1
T	**F**	**T**	
T	**F**	**F**	
F	**T**	**T**	
F	**T**	**F**	w_2
F	**F**	**T**	w_3
F	**F**	**F**	w_4

TABLE C.2. Schematic truth-table depicting the four (4) possible worlds compatible with the precondition of (A_3).

By (c) and (LO), p and q are not logically equivalent. Moreover, it is easy to verify that (f) will *not* be inaccurate in *any* of these four worlds, while (d) *must be inaccurate in world* w_2 (see the companion *Mathematica* notebook for these calculations). This suffices to show that $\succeq'$ *weakly* $\mathcal{I}_{\succeq}$-dominates $\succeq$.

VERIFYING THEOREM 3.4 AND THEOREM 3.5 [(WADA$_{\succeq}$) $\not\Rightarrow$ (A_5) and (WADA$_{\succeq}$) $\not\Rightarrow$ ($A_5^\star$)]

Suppose (a) $\langle p, q\rangle$ and $\langle p, r\rangle$ are mutually exclusive, (b) $q \succ r$, and (c) $p \vee r \succ p \vee q$. It can be shown that *there is no binary relation* $\succeq'$ on the agenda $\langle p, q, r\rangle$ such that (i) $\succeq'$ agrees with $\succeq$ on all judgments *except* (b) and (c), and (ii) $\succeq'$ weakly $\mathcal{I}_{\succeq}$-dominates $\succeq$. There are only four alternative judgment sets that need to be compared with $\{(b), (c)\}$, in terms of their $\mathcal{I}_{\succeq}$-values across the five possible worlds (w_1–w_5) compatible with the precondition of (A_5), which are depicted in Table C.3. (1) $\{q \sim r, p \vee r \succ p \vee q\}$, (2) $\{r \succ q, p \vee r \succ p \vee q\}$, (3) $\{q \succ r, p \vee r \sim p \vee q\}$, and

p	q	r	possible world w_i
T	**T**	**T**	
T	**T**	**F**	
T	**F**	**T**	
T	**F**	**F**	w_1
F	**T**	**T**	w_2
F	**T**	**F**	w_3
F	**F**	**T**	w_4
F	**F**	**F**	w_5

TABLE C.3. Schematic truth-table depicting the five (5) possible worlds compatible with the precondition of (A_5) [and ($A_5^\star$)].

(4) $\{q \sim r, p \vee r \sim p \vee q\}$. It is easy to verify that none of these alternative judgment sets weakly $\mathcal{I}_{\succeq}$-dominates the set $\{(b), (c)\}$, across the five salient possible worlds (see the companion *Mathematica* notebook for these calculations). Note: this argument establishes the *stronger* claim (Theorem 3.5) that (WADA) does *not* entail ($\mathrm{A}_5^{\star}$).

PROOF OF THEOREM 3.6
[($\mathcal{R}_{\succeq}$) ⇒ (WADA$_{\succeq}$) and ($\mathfrak{C}_4$) ⇒ (SADA$_{\succeq}$)]

We will first show that ($\mathfrak{C}_4$) ⇒ (SADA$_{\succeq}$) and then we'll prove ($\mathcal{R}_{\succeq}$) ⇒ (WADA$_{\succeq}$). Suppose ($\mathfrak{C}_4$) holds. That is, suppose $\Pr(\cdot)$ is a probability function that fully represents $\succeq$ (on agenda $\mathcal{A}$). Consider the expected $\mathcal{I}_{\succeq}$-inaccuracy, as calculated by $\Pr(\cdot)$, of $\succeq$. This is given by the sum over all possible worlds w of $\Pr(w) \cdot \mathcal{I}_{\succeq}(\succeq, w)$. Since $\mathcal{I}_{\succeq}(\succeq, w)$ is a sum of the components $\mathfrak{i}_{\succeq}(p \succeq q, w)$ for each pair of propositions $p, q \in \mathcal{A}$, and since expectations are linear, the expected inaccuracy is the sum of the expectation of these components. The expectation of the component $\mathfrak{i}_{\succeq}(p \succ q, w)$ is $2 \cdot \Pr(q \,\&\, \neg p)$ while the expectation of the component $\mathfrak{i}_{\succeq}(p \sim q, w)$ is $\Pr(p \not\equiv p)$. Now, there are two possibilities for a given pair $p, q \in \mathcal{A}$.

(1) $\Pr(p) > \Pr(q)$. In this case, the expected inaccuracy of $p \succ q$ is $2 \cdot \Pr(q \,\&\, \neg p)$, which is *strictly less than* the expected inaccuracy of either $p \sim q$, which is given by $\Pr(p \not\equiv p)$ or $q \succ p$, which is given by $2 \cdot \Pr(p \,\&\, \neg q)$. So, if $\Pr(p) > \Pr(q)$, then $p \succ q$ *uniquely* minimizes expected $\mathfrak{i}_{\succeq}$-inaccuracy.

(2) $\Pr(p) = \Pr(q)$. In this case, the expected inaccuracy of $p \sim q$ is $\Pr(p \not\equiv p)$, which *is equal to* the expected inaccuracy of $p \succ q$, which is given by $2 \cdot \Pr(\neg p \,\&\, q)$. So, if $\Pr(p) = \Pr(q)$, then $p \sim q$ non-uniquely minimizes expected $\mathfrak{i}_{\succeq}$-inaccuracy, *i.e.*, all judgments have the same $\mathfrak{i}_{\succeq}$-expectation.

Thus, since $\Pr(\cdot)$ fully represents $\succeq$, this means that, $\succeq$ will have an expected $\mathcal{I}_{\succeq}$-inaccuracy that is *no greater than* that of any other $\succeq$-relation (on $\mathcal{A}$) with respect to $\Pr(\cdot)$. Thus, the measure $\mathcal{I}_{\succeq}$ is *evidentially proper*. And, no relation $\succeq'$ can *strictly* dominate $\succeq$ in $\mathcal{I}_{\succeq}$-inaccuracy (on $\mathcal{A}$), since such a $\succeq'$ would have a *strictly lower* Pr-expected $\mathcal{I}_{\succeq}$-inaccuracy than $\succeq$ (on $\mathcal{A}$). In other words, ($\mathfrak{C}_4$) entails (SADA$_{\succeq}$). Now, suppose that ($\mathcal{R}_{\succeq}$) holds. That is, suppose $\Pr(\cdot)$ is a *regular* probability function that fully represents $\succeq$ (on $\mathcal{A}$). Then, no relation $\succeq'$ can *weakly* dominate $\succeq$ in $\mathcal{I}_{\succeq}$-inaccuracy (on $\mathcal{A}$), since such a relation would have a *strictly lower* Pr-expected $\mathcal{I}_{\succeq}$-inaccuracy than $\succeq$ (on $\mathcal{A}$). Therefore, ($\mathcal{R}_{\succeq}$) entails (WADA$_{\succeq}$).

VERIFYING THE UNIQUENESS OF OUR EVIDENTIALLY PROPER MEASURE $\mathfrak{i}_{\succeq}$

THEOREM. *$a := 2$ and $b := 0$ is the only numerical assignment to a and b which ensures that the following parametric definition of $\mathfrak{i}_{\succeq}$ is evidentially proper.*

$$\mathfrak{i}_{\succeq}(p \succeq q, w) \overset{\text{def}}{=} \begin{cases} a & \text{if } q \,\&\, \neg p \text{ is true in } w, \text{ and } p \succ q, \\ b & \text{if } q \equiv p \text{ is true in } w, \text{ and } p \succ q, \\ 1 & \text{if } p \not\equiv q \text{ is true in } w, \text{ and } p \sim q, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. Let $\mathfrak{m}_4 = \Pr(p \,\&\, q)$, $\mathfrak{m}_3 = \Pr(\neg p \,\&\, q)$, and $\mathfrak{m}_2 = \Pr(p \,\&\, \neg q)$. Then, the following claim (which implicitly *universally quantifies* over these three probability masses $\mathfrak{m}_2, \mathfrak{m}_3, \mathfrak{m}_4$, and presupposes that they are each on $[0,1]$ and that they sum to at most 1) states that this parametric definition of $\acute{\imath}_{\succeq}$ is *evidentially proper*.

$$\mathfrak{m}_2 + \mathfrak{m}_4 > \mathfrak{m}_3 + \mathfrak{m}_4 \Rightarrow \left(\begin{array}{c} a \cdot \mathfrak{m}_3 + b \cdot (1 - (\mathfrak{m}_2 + \mathfrak{m}_3)) \leq a \cdot \mathfrak{m}_2 + b \cdot (1 - (\mathfrak{m}_2 + \mathfrak{m}_3)) \\ \& \\ a \cdot \mathfrak{m}_3 + b \cdot (1 - (\mathfrak{m}_2 + \mathfrak{m}_3)) \leq \mathfrak{m}_2 + \mathfrak{m}_3 \end{array} \right)$$

$$\&$$

$$\mathfrak{m}_2 + \mathfrak{m}_4 = \mathfrak{m}_3 + \mathfrak{m}_4 \Rightarrow \left(\begin{array}{c} \mathfrak{m}_2 + \mathfrak{m}_3 \leq a \cdot \mathfrak{m}_2 + b \cdot (1 - (\mathfrak{m}_2 + \mathfrak{m}_3)) \\ \& \\ \mathfrak{m}_2 + \mathfrak{m}_3 \leq a \cdot \mathfrak{m}_3 + b \cdot (1 - (\mathfrak{m}_2 + \mathfrak{m}_3)) \end{array} \right)$$

There is a *unique* numerical assignment to the parameters a and b which makes this universal claim true, and it is $a := 2; b := 0$. This implies that our scoring rule is (essentially) the *unique* evidentially proper scoring rule for comparative confidence judgments. It also implies that there is no (truth-directed — *fn.* 30) proper scoring rule for the weaker notion of inaccuracy associated with the assumption of a unique vindicated ordering $\mathring{\succeq}_w$ (*fn.* 74).[122] We discovered this result using *Mathematica*'s quantifier elimination algorithm for the theory of real closed fields (see the companion *Mathematica* notebook for the verification of this theorem). □

VERIFYING THEOREM 4.1
[$(\mathcal{R}_{\succeq}) \not\Rightarrow (\text{WADA}_{\succeq})$ — for the *Kemeny measure* of inaccuracy for $\succeq$-relations]

Consider a Boolean algebra $\mathcal{B}_8$ generated by three states $\{\mathfrak{s}_1, \mathfrak{s}_2, \mathfrak{s}_3\}$. And, consider the comparative confidence relation $\succeq$ on $\mathcal{B}_8$, the adjacency matrix of which is depicted in Table C.4.

$\succeq$	$\bot$	$\mathfrak{s}_1$	$\mathfrak{s}_2$	$\mathfrak{s}_3$	$\mathfrak{s}_1 \vee \mathfrak{s}_2$	$\mathfrak{s}_1 \vee \mathfrak{s}_3$	$\mathfrak{s}_2 \vee \mathfrak{s}_3$	$\top$
$\bot$	1	0	0	0	0	0	0	0
$\mathfrak{s}_1$	1	1	0	0	0	0	0	0
$\mathfrak{s}_2$	1	1	1	0	0	0	0	0
$\mathfrak{s}_3$	1	1	1	1	1	0	0	0
$\mathfrak{s}_1 \vee \mathfrak{s}_2$	1	1	1	0	1	0	0	0
$\mathfrak{s}_1 \vee \mathfrak{s}_3$	1	1	1	1	1	1	0	0
$\mathfrak{s}_2 \vee \mathfrak{s}_3$	1	1	1	1	1	1	1	0
$\top$	1	1	1	1	1	1	1	1

TABLE C.4. Adjacency matrix of a fully Pr-representable relation $\succeq$ that is weakly dominated in Kemeny distance from vindication.

[122] Similar negative results regarding the non-existence of proper scoring rules have recently been proven for imprecise probabilities (Seidenfeld et al., 2012).

This relation $\succeq$ is fully represented by the regular probability function $\Pr_{\mathfrak{m}}(\cdot)$ determined by the following probability mass function:

$$\mathfrak{m}_1 = 1/16.$$
$$\mathfrak{m}_2 = 5/16.$$
$$\mathfrak{m}_3 = 5/8.$$

The relation $\succeq$ is weakly dominated in Kemeny distance from vindication by the relation $\succeq'$, the adjacency matrix of which is depicted in Table C.5.

$\succeq'$	$\bot$	$\mathfrak{s}_1$	$\mathfrak{s}_2$	$\mathfrak{s}_3$	$\mathfrak{s}_1 \vee \mathfrak{s}_2$	$\mathfrak{s}_1 \vee \mathfrak{s}_3$	$\mathfrak{s}_2 \vee \mathfrak{s}_3$	$\top$
$\bot$	1	0	0	0	0	0	0	0
$\mathfrak{s}_1$	1	1	1	0	0	0	0	0
$\mathfrak{s}_2$	1	1	1	0	0	0	0	0
$\mathfrak{s}_3$	1	1	1	1	0	0	0	0
$\mathfrak{s}_1 \vee \mathfrak{s}_2$	1	1	1	1	1	0	0	0
$\mathfrak{s}_1 \vee \mathfrak{s}_3$	1	1	1	1	1	1	1	0
$\mathfrak{s}_2 \vee \mathfrak{s}_3$	1	1	1	1	1	1	1	0
$\top$	1	1	1	1	1	1	1	1

TABLE C.5. Adjacency matrix of a fully Pr-representable $\succeq'$ that weakly dominates $\succeq$ (Table C.4) in Kemeny distance from vindication.

The relation $\succeq'$ is fully represented by the regular probability function $\Pr_{\mathfrak{m}'}(\cdot)$ determined by the following probability mass function:

$$\mathfrak{m}'_1 = 7/24.$$
$$\mathfrak{m}'_2 = 7/24.$$
$$\mathfrak{m}'_3 = 5/12.$$

Let w_i be the possible world corresponding to state $\mathfrak{s}_i$ of $\mathcal{B}_8$. It is then straightforward to verify that $\succeq'$ weakly dominates $\succeq$ in Kemeny inaccuracy (*viz.*, δ-inaccuracy), *i.e.*, that the following claims are true (see the companion *Mathematica* notebook).

(*i*) $(\forall w_i)\,[\delta(\succeq', w_i) \leq \delta(\succeq, w_i)]$, and

(*ii*) $(\exists w_i)\,[\delta(\succeq', w_i) < \delta(\succeq, w_i)]$.

Therefore, because probabilistic representability of $\succeq$ (even by a *regular* probability function) is not sufficient to ensure that $\succeq$ is not weakly dominated in Kemeny inaccuracy, the Kemeny measure of inaccuracy δ is not evidentially proper (even for regular evidential probability functions). This can be seen by calculating the expected Kemeny inaccuracy of $\succeq$ and $\succeq'$ above, relative to the evidential representer of $\succeq$ ($\Pr_{\mathfrak{m}}$). In the companion *Mathematica* notebook, we perform these calculations, which reveal that $\succeq'$ has a lower expected Kemeny inaccuracy than $\succeq$, relative to $\Pr_{\mathfrak{m}}$. Of course, this is not surprising, since it would be true for *any* regular probability function. □

APPENDIX D

Theorems from Part III

PROOF OF THEOREM 6.1
[If b' dominates b (in Brier score), and both take values only in $[0, 1]$, then *either* $b(P) > b'(P)$ and $b(\neg P) > b'(\neg P)$ *or* $b(P) < b'(P)$ and $b(\neg P) < b'(\neg P)$.]

PROOF. If b' dominates b, then b' must have lower inaccuracy both in w_1 and in w_2. The inaccuracy of a credence function in w_1 is the sum of two terms. If $b'(P) > b(P)$, then the first term is greater for b' than for b. Thus, if b' is less inaccurate than b in w_1, then the second term must be *lower* for b' than for b. But this means that $b'(\neg P) > b(\neg P)$. By considering the two terms summing to the inaccuracy in w_2, we can show that if $b'(P) < b(P)$, then $b'(\neg P) < b(\neg P)$. Thus, if $b'(P) \neq b(P)$, then $b'(\neg P)$ must also differ from $b(\neg P)$, in the same direction. Similar reasoning establishes the converse. $\square$

VERIFYING THEOREM 6.2
[For any coherent function b' that Brier-dominates S's credence function b with respect to ϕ and ψ, there exist quantities α and β that are symmetrically inter-definable with respect to ϕ and ψ, *via* the following symmetric inter-translations.

$$\alpha = \frac{1}{2}\phi + \frac{1}{2}\psi + \frac{1}{16}\left(\frac{\phi+\psi}{\phi-\psi}\right) \qquad \beta = \frac{1}{2}\phi + \frac{1}{2}\psi - \frac{1}{16}\left(\frac{\phi+\psi}{\phi-\psi}\right)$$

$$\phi = \frac{1}{2}\alpha + \frac{1}{2}\beta + \frac{1}{16}\left(\frac{\alpha+\beta}{\alpha-\beta}\right) \qquad \psi = \frac{1}{2}\alpha + \frac{1}{2}\beta - \frac{1}{16}\left(\frac{\alpha+\beta}{\alpha-\beta}\right)$$

Where b *Brier-dominates* b' with respect to α and β.]

See the companion *Mathematica* notebook for a verification of this theorem.

Bibliography

Barrett, M. (2010). The possibility of infinitesimal chances. In *The Place of Probability in Science*, pp. 65–79. Springer.

Belnap, N. (1977). A useful four-valued logic. In *Modern uses of multiple-valued logic*, pp. 5–37.

Berker, S. (2013). Epistemic teleology and the separateness of propositions. *Philosophical Review 122*, 337–93.

Boghossian, P. (2003). The normativity of content. *Philosophical Issues 13*(1), 31–45.

BonJour, L. (1985). *The structure of empirical knowledge*. Cambridge University Press.

Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review 78*(1), 1–3.

Briggs, R. (2009). Distorted reflection. *Philosophical Review 118*(1), 59–85.

Briggs, R., F. Cariani, K. Easwaran, and B. Fitelson (2014). Individual coherence and group coherence. In J. Lackey (Ed.), *Essays in Collective Epistemology*. OUP. To appear.

Bronfman, A. (ms.). A gap in joyce's argument for probabilism. Manuscript.

Broome, J. (2007). Wide or narrow scope? *Mind 116*(462), 359–370.

Buchak, L. (2013). Belief, credence, and norms. *Philosophical Studies*, 1–27.

Caie, M. (2013). Rational probabilistic incoherence. To appear in *The Philosophical Review*.

Campbell-Moore, C. (2014a). How to express self-referential probability and avoid the (bad) consequences. Manuscript.

Campbell-Moore, C. (2014b). Rational probabilistic incoherence? a reply to michael caie. Manuscript.

Capotorti, A. and B. Vantaggi (2000). Axiomatic characterization of partial ordinal relations. *International Journal of Approximate Reasoning 24*(2), 207–219.

Carnap, R. (1962). Logical foundations of probability.

Carr, J. (2013). What to expect when you're expecting. Manuscript.

Christensen, D. (1996). Dutch-book arguments depragmatized: Epistemic consistency for partial believers. *The Journal of Philosophy 93*(9), 450–479.

Christensen, D. (2004). *Putting logic in its place*. Oxford University Press.

Clifford, W. (1877). The ethics of belief. *Contemporary Review 29*, 289–309.

Conee, E. and R. Feldman (2004). *Evidentialism: essays in epistemology*. Oxford University Press.

Csiszár, I. (1991). Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *Ann. Statist. 19*(4), 2032–2066.

de Finetti, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré 7*, 1–68.

de Finetti, B. (1951). *La 'Logica del Plausible' secondo la concezione di Polya*. Società Italiana per il Progresso delle Scienze.

de Finetti, B. (1970). *Theory of probability*. Wiley.

Dempster, A. P. (1968). A generalization of Bayesian inference. (With discussion). *J. Roy. Statist. Soc. Ser. B 30*, 205–247.

DePaul, M. (2001). Value monism in epistemology. In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford University Press.

Deza, M. and E. Deza (2009). *Encyclopedia of distances*. Springer.

Dretske, F. (2013). Gettier and justified true belief: fifty years on. *The Philosophers' Magazine 61*. URL: `http://philosophypress.co.uk/?p=1171`.

Duddy, C. and A. Piggins (2012). A measure of distance between judgment sets. *Social Choice and Welfare 39*(4), 855–867.

Easwaran, K. (2013a). Decision theory without representation theorems. Manuscript.

Easwaran, K. (2013b). Dr. Truthlove, or how I learned to stop worrying and love Bayesian probabilities. Manuscript.

Easwaran, K. (2014). Regularity and hyperreal credences. *Philosophical Review 123*(1), 1–41.

Easwaran, K. and B. Fitelson (2012). An "evidentialist" worry about Joyce's argument for probabilism. *Dialectica 66*(3), 425–433.

Easwaran, K. and B. Fitelson (2014). Accuracy, coherence and evidence. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology, Vol. 5*. Oxford University Press. To appear.

Field, H. (2008). *Saving truth from paradox*. Oxford University Press Oxford.

Fine, K. (1975). Vagueness, truth and logic. *Synthese 30*(3), 265–300.

Fine, T. (1973). *Theories of probability*. Academic Press.

Fishburn, P. C. (1986). The axioms of subjective probability. *Statistical Science 1*(3), 335–345.

Fitelson, B. (2008). A decision procedure for probability calculus with applications. *The Review of Symbolic Logic 1*(01), 111–125.

Fitelson, B. (2012). Accuracy, language dependence, and joyce's argument for probabilism. *Philosophy of Science 79*(1), 167–174.

Foley, R. (1992). *Working without a net*. Oxford University Press.

Foley, R. (2009). Beliefs, degrees of belief, and the lockean thesis. In *Degrees of belief*, pp. 37–47. Springer.

Forrest, P. (1989). The problem of representing incompletely ordered doxastic systems. *Synthese 79*(2), 279–303.

Friedman, J. (2013). Suspended judgment. *Philosophical studies 162*(2), 165–181.

Frisch, M. (2005). *Inconsistency, asymmetry, and non-locality: A philosophical investigation of classical electrodynamics*. Oxford University Press.

Fumerton, R. (1995). *Metaepistemology and skepticism*. Rowman & Littlefield.

Gibbard, A. (2005). Truth and correct belief. *Philosophical Issues 15*(1), 338–350.

Goldman, A. (1999). *Knowledge in a social world*. Oxford University Press.

Greaves, H. (2013). Epistemic decision theory. To appear in *Mind*.

Hájek, A. (2008). Arguments for–or against–probabilism? *The British Journal for the Philosophy of Science 59*(4), 793–819.

Hájek, A. (2010). A puzzle about degree of belief. Manuscript.

Hájek, A. (2012). Staying regular? Manuscript.

Halpern, J. Y. (2003). *Reasoning about uncertainty*. MIT Press.

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

Hawthorne, J. (2009). The lockean thesis and the logic of belief. In *Degrees of Belief*, pp. 49–74. Springer.

Hawthorne, J. and L. Bovens (1999). The preface, the lottery, and the logic of belief. *Mind 108*(430), 241–264.

Hedden, B. (2013). Time-slice rationality. Manuscript.

Hempel, C. (1962). Deductive-nomological vs statistical explanation. *Minnesota studies in the philosophy of science 3*, 98–169.

Humberstone, L. (1992). Direction of fit. *Mind*, 59–83.

Icard, T. (2014). Decision-theoretic considerations on comparative probability. Manuscript.

James, W. (1896). The will to believe. *The New World 5*, 327–347.

Jeffrey, R. (1970). Dracula meets wolfman: Acceptance vs. partial belief. In *Induction, acceptance and rational belief*, pp. 157–185. Springer.

Jeffrey, R. (1992). *Probability and the Art of Judgment*. Cambridge University Press.

Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Phil. Sci. 65*(4), 575–603.

Joyce, J. (2005). How probabilities reflect evidence. *Phil. Perspectives 19*(1), 153–178.

Joyce, J. (2009). Accuracy and coherence: prospects for an alethic epistemology of partial belief. In F. Huber and C. Schmidt-Petri (Eds.), *Degrees of Belief*. Springer.

Joyce, J. (2013). The role of evidence in an accuracy-centered epistemology. Manuscript.

Kaplan, M. (2013). Coming to terms with our human fallibility: Christensen on the preface. *Philosophy and Phenomenological Research 87*(1), 1–35.

Keynes, J. M. (1921). *A treatise on probability*. Macmillan.

Klein, P. (1985). The virtues of inconsistency. *The Monist 68*(1), 105–135.

Kolodny, N. (2007). How does coherence matter? In *Proceedings of the Aristotelian Society*, Volume 107, pp. 229–263.

Kolodny, N. (2008). Why be disposed to be coherent? *Ethics 118*(3), 437–463.

Kolodny, N. and J. MacFarlane (2010). Ifs and oughts. *Journal of Philosophy 107*, 115–143.

Konek, J. (2014). Vindicating scott's laws of comparative probability using non-additive scoring rules. Lecture notes.

Konek, J. and B. Levinstein (2014). The foundations of epistemic decision theory. Manuscript.

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

Korb, K. (1992). The collapse of collective defeat: Lessons from the lottery paradox. In *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pp. 230–236.

Kraft, C. H., J. W. Pratt, and A. Seidenberg (1959). Intuitive probability on finite sets. *The Annals of Mathematical Statistics 30*(2), 408–419.

Kyburg, H. (1970). Conjunctivitis. In M. Swain (Ed.), *Induction, acceptance and rational belief*, pp. 55–82. Reidel.

Lehrer, K. and C. Wagner (1985). Intransitive indifference: The semi-order problem. *Synthese 65*(2), 249–256.

Leitgeb, H. (2013). The stability theory of belief. Manuscript.

Leitgeb, H. (2014). A way out of the preface paradox? *Analysis 74*(1), 11–15.

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

Levinstein, B. (2013). *Accuracy as Epistemic Utility*. Ph. D. thesis, Rutgers University.

Lewis, D. (1981). A subjectivist's guide to objective chance. In *Ifs*, pp. 267–297. Springer.

Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy 61*(4), 343–377.

Lewis, D. (1994). Humean supervenience debugged. *Mind*, 473–490.

Littlejohn, C. (2012). *Justification and the truth-connection.* Cambridge University Press.

Liu, L. and R. Yager (Eds.) (2008). *Classic works of the Dempster-Shafer theory of belief functions*, Volume 219 of *Studies in Fuzziness and Soft Computing.* Springer, Berlin.

MacFarlane, J. (2004). In what sense (if any) is logic normative for thought? Manuscript.

Maher, P. (2002). Joyce's argument for probabilism. *Philosophy of Science 69*(1), 73–81.

Merricks, T. (1995). Warrant entails truth. *Philos. and Phenom. Research 55*(4), 841–855.

Miller, D. (1974). Popper's qualitative theory of verisimilitude. *British Journal for the Philosophy of Science*, 166–177.

Miller, D. (1975). The accuracy of predictions. *Synthese 30*(1), 159–191.

Miller, D. (2006). *Out of error: Further essays on critical rationalism.* Aldershot: Ashgate Publishing Company.

Moss, S. (2013). Time-slice epistemology and action under indeterminacy. Manuscript.

Nelkin, D. (2000). The lottery paradox, knowledge, and rationality. *The Philosophical Review 109*(3), 373–409.

Parfit, D. (1988). What we together do. Manuscript.

Pedersen, A. P. and C. Glymour (2012). What language dependence problem? a reply for joyce to fitelson on joyce. *Philosophy of Science 79*(4), 561–574.

Pettigrew, R. (2011). Epistemic utility arguments for probabilism. *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition),
URL = <`http://plato.stanford.edu/entries/epistemic-utility/`>.

Pettigrew, R. (2013a). Accuracy and evidence. To appear in *Dialectica*.

Pettigrew, R. (2013b). Epistemic utility and norms for credences. *Philosophy Compass*, URL = <`http://dx.doi.org/10.1111/phc3.12079`>.

Pettigrew, R. (2014). How should we measure accuracy in epistemology? a new result. <`http://m-phi.blogspot.com/2014/04/how-should-we-measure-accuracy-in.html`>.

Pollock, J. (1983). Epistemology and probability. *Synthese 55*(2), 231–252.

Pollock, J. (1986). The paradox of the preface. *Philosophy of Science*, 246–258.

Pollock, J. (1990). *Nomic probability and the foundations of induction.* Oxford Univ. Press.

Pollock, J. (1995). *Cognitive carpentry: A blueprint for how to build a person.* Mit Press.

Popper, K. (1972). *Objective Knowledge: An Evolutionary Approach* (Second ed.). New York: Oxford University Press.

Predd, J. B., R. Seiringer, E. H. Lieb, D. N. Osherson, H. V. Poor, and S. R. Kulkarni (2009). Probabilistic coherence and proper scoring rules. *IEEE Transactions on Information Theory 55*(10), 4786–4792.

Priest, G. (2002). *Beyond the limits of thought.* Oxford University Press.

Pruss, A. (2012). Infinite lotteries, perfectly thin darts and infinitesimals. *Thought: A Journal of Philosophy 1*(2), 81–89.

Pruss, A. (2013). Infinitesimals are too small for countably infinite fair lotteries. *Synthese*, 1–7.

Ramsey, F. (1928). Truth and probability. In R. Braithwaite (Ed.), *The foundations of mathematics and other logical essays (1931)*, pp. 156–198. Routledge.

Rosenkrantz, R. (1981). Foundations and applications of inductive probability.

Ross, J. and M. Schroeder (2012). Belief, credence, and pragmatic encroachment1. *Philosophy and Phenomenological Research*.

Ryan, S. (1991). The preface paradox. *Philosophical studies 64*(3), 293–307.

Ryan, S. (1996). The epistemic virtues of consistency. *Synthese 109*(2), 121–141.

Savage, L. (1972). *The foundations of statistics.* Dover.

Schervish, M., T. Seidenfeld, and J. Kadane (2009). Proper scoring rules, dominated forecasts, and coherence. *Decision Analysis 6*(4), 202–221.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2012). Forecasting with imprecise probabilities. Manuscript.

Shafer, G. (1976). *A mathematical theory of evidence.* Princeton university press.

Shah, N. (2003). How truth governs belief. *The Philosophical Review 112*(4), 447–482.

Shimony, A. (1955). Coherence and the axioms of confirmation. *Journal of Symbolic Logic*, 1–28.

Smith, M. (2005). Meta-ethics. *The Oxford Handbook of Contemporary Philosophy*, 3–30.

Sosa, E. (2003). The place of truth in epistemology. In *Intellectual Virtue: Perspectives From Ethics and Epistemology*, pp. 155–180. New York: Oxford University Press.

Stalnaker, R. (1991). The problem of logical omniscience, I. *Synthese 89*(3), 425–440.

Stalnaker, R. (1999). The problem of logical omniscience II. In *Context and content: Essays on intentionality in speech and thought*, pp. 255–273. Oxford University Press.

Steinberger, F. (2014). Explosion and the normativity of logic. To appear in *Mind*.

Sturgeon, S. (2008). Reason and the grain of belief. *Noûs 42*(1), 139–165.

Thomson, J. J. (2008). *Normativity.* Open Court.

Tichỳ, P. (1974). On popper's definitions of verisimilitude. *The British Journal for the Philosophy of Science 25*(2), 155–160.

Titelbaum, M. (2013). *Quitting certainties: a Bayesian framework modeling degrees of belief.* Oxford University Press.

Wedgwood, R. (2002). The aim of belief. *Philosophical Perspectives 16*, 267–297.

Williams, R. (2012). Gradational accuracy and non-classical semantics. *Review of Symbolic Logic 5*(4), 513–537.

Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press.

Williamson, T. (2007). How probable is an infinite sequence of heads? *Analysis 67*(3), 173–180.

Wong, S. M., Y. Yao, P. Bollmann, and H. Burger (1991). Axiomatization of qualitative belief structure. *IEEE Transactions on Systems, Man and Cybernetics 21*(4), 726–734.

Zagzebski, L. (1994). The inescapability of Gettier problems. *Phil. Quarterly 44*(174), 65–73.

Zardini, E. (2011). Truth without contra (di) ction. *The Review of Symbolic Logic 4*(04), 498–535.