

## What Are Degrees of Belief?

**Abstract.** Probabilism is committed to two theses:

- 1) Opinion comes in degrees—call them *degrees of belief*, or *credences*.
- 2) The degrees of belief of a rational agent obey the probability calculus.

Correspondingly, a natural way to argue for probabilism is:

- i) to give an account of what degrees of belief *are*,

and then

- ii) to show that *those things* should be probabilities, on pain of irrationality.

Most of the action in the literature concerns stage ii). Assuming that stage i) has been adequately discharged, various authors move on to stage ii) with varied and ingenious arguments. But an unsatisfactory response at stage i) clearly undermines any gains that might be accrued at stage ii) as far as probabilism is concerned: if *those things* are not degrees of belief, then it is irrelevant to probabilism whether *they* should be probabilities or not.

In this paper we scrutinize the state of play regarding stage i). We critically examine several of the leading accounts of degrees of belief: reducing them to corresponding betting behavior (de Finetti); measuring them by that behavior (Jeffrey); and analyzing them in terms of preferences and their role in decision-making more generally (Ramsey, Lewis, Maher). We argue that the accounts fail, and so they are unfit to subserve arguments for probabilism. We conclude more positively: ‘degree of belief’ should be taken as a primitive concept that forms the basis of our best theory of rational belief and decision: probabilism.

*Keywords:* probabilism, rationality, degrees of belief, betting interpretation, Dutch book argument, representation theorem, radical interpretation, primitivism.

Probabilism is committed to two theses:

- 1) Opinion comes in degrees—call them *degrees of belief*, or *credences*.
- 2) The degrees of belief of a rational agent obey the probability calculus.

Correspondingly, a natural way to argue for probabilism is:

- i) to give an account of what degrees of belief *are*,

and then

- ii) to show that *those things* should be probabilities, on pain of irrationality.

Most of the action in the literature concerns stage ii). Assuming that stage i) has been adequately discharged, various authors move on to stage ii) with varied and ingenious arguments. But an unsatisfactory response at stage i) clearly undermines any gains that might be accrued at stage ii) as far as probabilism is concerned: if *those things* are not degrees of belief, then it is irrelevant to probabilism whether *they* should be probabilities or not.

We would like to take a step back and scrutinize the state of play regarding stage i). Indeed, we find it surprising that it hasn't been given more scrutiny than it has. Surprising, because 'degree of belief' is *the* central notion in subjective probability theory, the core concept around which the entire theory is built. 'Degree of belief' is supposed to underwrite a more plausible and nuanced probabilist epistemology than the traditional 'Cartesian' epistemology that traffics only in binary, all-or-nothing notions such as 'knowledge' and 'belief'. Meanwhile, 'degree of belief' is also offered as an interpretation of the probability calculus, sitting proudly alongside the classical, frequentist, logical, and propensity interpretations. Thus, tunnels with opposite starting points and heading in opposite directions are intended to meet happily in the middle: the graded notion of 'degree of belief' supposedly obeys the probability calculus, and that formalism supposedly finds a familiar interpretation in 'degree of belief'. And never mind our theoretical commitments. It seems to be a platitude that there are such things as degrees of belief—just ask any punter at the racetrack. Isn't it our job as philosophers to try to understand what they are?

So far we have failed. One sees, for example, frequent gestures at some kind of betting interpretation, often accompanied with a slightly coy acknowledgment that this interpretation, well, isn't *strictly speaking* correct. But one is then immediately assuaged: at least it's "approximately correct", or "correct over a significant range of cases", or a "useful idealization", or what have you. Like the DN theory of scientific explanation, and Jason in the *Friday the 13<sup>th</sup>* horror movie series, the interpretation seems never to die. Having thus glossed over the numerous problems with that interpretation—of which, more soon—the stage is then set for the *pièce de résistance*, the argument for 2). But from the fact that betting prices putatively should obey the probability calculus, it does not follow that degrees of belief, which are not extensionally equivalent to them, should do so also.

More generally, every account of 'degree of belief' that we will discuss seeks to connect it to *preferences*, in the service of an eventual argument for probabilism. The arguments supposedly prove that some preference-based thing obeys the probability calculus. The question remains, however: does that thing deserve the name 'degree of belief'? In fact, every account fails to

establish even a contingent connection between the two, still less a necessary connection, still less a sound argument for probabilism.

However, we do not want merely to be purveyors of doom and gloom. Although much of our discussion will be negative, we will conclude more positively: we should not be afraid of taking ‘degree of belief’ as a primitive in our conceptual apparatus. We can afford to do so, because we have so many ways of triangulating to its nature, thanks in large part to all the important work that has previously gone into the various analyses that we ultimately reject. Taking a concept to be primitive—i.e. unanalyzable, or at least unanalyzed, in terms of other, more basic concepts—does not preclude us from saying many illuminating things about it. This, too, ought to be a platitude, once we reflect on our treatment of other primitive concepts, such as length, time, or mass. Each of them resists further reduction, but that hardly renders us unable to say substantive things about them. Much of analytic philosophy is a series of failed attempts to provide analyses of our core concepts (knowledge, causation, personal identity, morally wrong action, . . . ), yet much light has been shed on these concepts in the process. And as we will see, one substantive thing we can say about degrees of belief is that they should obey the probability calculus.

In the sequel, we attempt to map out several of the leading accounts of degrees of belief and show why they cannot be the basis of an argument for probabilism: the betting interpretation (in both its ‘actual’ and ‘hypothetical’ variants), the view that betting prices merely measure credences, and the representation theorem approach pursued in three slightly different ways. Partly because of lack of space, and partly because textual exegesis and historical scholarship are neither our main strength nor main concern, we do not pretend to chart every change in view that some author may have displayed in this paper or that. If, for example, a late article of de Finetti’s in Italian shows that he recanted some view that we attribute to him, it does not matter for our purposes. Each position that we will consider deserves its day in the sun, even if it was only held by a time-slice of one author or other (and in a couple of cases, perhaps not even that). We wish to evaluate the positions themselves. We conclude with primitivism, the position that at least our current time-slices endorse.

### **Actual operationalism: de Finetti**

It is only appropriate to begin with one of the pioneers of the subjective interpretation of probability, de Finetti. His earlier writings coincided with the heyday of operationalism and positivism. These movements predated

the heyday of behaviorism, but there are obvious parallels there. Actually, de Finetti's analysis of 'degree of belief' can be regarded as a particularly localized behaviorism: mental states are not only analysed in terms of behaviors and behavioral dispositions, but very specific ones at that.

Here is an admirably clear—one might say unflinching—statement of the operationalist position by de Finetti (1990):

In order to give an effective meaning to a notion and not merely an appearance of such in a metaphysical-verbalistic sense an operational definition is required. By this we mean a definition based on a criterion which allows us to measure it. (76)

And his definition of subjective probability meets his own strictures:

The probability  $P(E)$  that You attribute to an event  $E$  is therefore the certain gain  $p$  which You judge equivalent to a unit gain conditional on the occurrence of  $E$ : in order to express it in a dimensionally correct way, it is preferable to take  $pS$  equivalent to  $S$  conditional on  $E$ , where  $S$  is any amount whatsoever, one Lira or one million, \$20 or £75. (75)

This is the famous *betting interpretation* of degrees of belief.

Monetary bets make attractive primitives. Like love, longevity, and Sydney beachfront property, money is almost universally desired. Moreover, money is usually readily available (unlike Sydney beachfront property), easily subdivided (unlike longevity), and easily quantified (unlike love). Note that the betting interpretation already imposes a certain structure on credences. For example, you can't pay the complex amount  $\$(0.3 + 2i)$  on a bet; credences so understood can't be *complex* numbers. Nor can you pay an infinitesimal amount for a bet; credences so understood can't be *infinitesimal* numbers, pace Lewis (1980) and Skyrms (1980). Less esoterically, Kolmogorov's axiomatization assumes that probabilities are *real* numbers between 0 and 1 inclusive, but even most of those numbers are ineligible to be credences according to the betting interpretation—after all, monetary amounts are only finitely divisible, so only *rational* probabilities can be given a betting interpretation, and indeed comparatively few of *them*. (Now is not the time to protest that pay-offs should be measured in utiles—we will consider them soon enough.) Moreover, credences must be *precise* (although this assumption can be relaxed if we allow the buying and selling price of a given bet to differ). To this extent, the very definition of credence here already places certain probabilistic constraints on it. We are thus well on the way towards an argument for probabilism.

Of course, there are numerous well-known objections to the betting interpretation. But somehow their very familiarity seems to encourage people to dismiss them: familiarity breeds contentment, perhaps. In fact, the arguments lose none of their potency for that. And as the arguments accumulate, reinforcing each other and successively closing off escape routes, one begins to doubt that the betting interpretation is even a good approximation, or correct over a significant range of cases, or a useful idealization, or what have you. This will be important when we look at the prospects for its descendants.

For starters, the betting interpretation suffers from general problems with operationalism. Compare the fate of giving an operational definition of ‘temperature’ in terms of what thermometers say. We surely want to allow that thermometers can be ill-calibrated—indeed, to insist that all actual thermometers are at least *a little* ill-calibrated. But it is a consequence of operationalism that necessarily they are perfectly calibrated, to infinite precision. And so it is with the analysis of credences as betting prices: it is not *possible* for them to differ, even in the 100th decimal place. See Joyce (1998) for further arguments against such operational definitions of degrees of belief.

More generally, the betting interpretation suffers from analogues of familiar problems with behaviorism—so much so that it is perhaps surprising how the former has much currency even in 2007, while the latter died in the 1950’s. And even qualified resurrections of the latter, which claim that it is at least “approximately correct”, or “correct over a significant range of cases”, or a “useful idealization”, are not faring well. Recall the *perfect actor* objection to behaviorism (attributed to Putnam). The betting interpretation makes it far too easy for us to be perfect actors with respect to our credences—think of the football team owner who flamboyantly places a bet at ridiculously short odds on his team winning, as a display of his loyalty. Recall Putnam’s (1965) *superspartan* objection, or Strawson’s (1994) *weather-watchers* objection to behaviorism, which imagine agents in whom the usual behavioral manifestations of mental states are absent. The betting interpretation makes it far too easy for us to be superspartans or weather-watchers with respect to our credences—think of the puritan who disdains all betting, and who is wrongly judged either to have no credences at all, or to have them uniformly set to 0 (the only price he will pay for any bet). Recall the *Chisholm-Geach* objection to behaviorism, which points out its circularity: one cannot specify the behavioral dispositions that some mental state will result in without adverting to other mental states (Chisholm 1957). The betting interpretation is apparently circular, because your betting price

regarding some proposition  $X$  is determined not just by your credence in  $X$ , but also by a raft of your credences in other propositions; a proper analysis of ‘your credence in  $X$  is  $p$ ’ must thus take the form ‘your betting price for  $X$  is  $p$ , *provided your credences in such-and-such are thus-and-so.*’<sup>1</sup> For example, you had better have credences about what bets are and what their prizes are, you had better not have defeating credences about any downstream consequences of having a betting price of  $p$  (e.g. some credence that God eternally punishes those who have that betting price), and so on. Of course, we can take these conditions for granted in real life; but an *analysis* cannot.

Still more generally, the betting interpretation suffers from some familiar problems with frequentism about probability, the Humean regularity theory of causation, and the naïve regularity theory of lawhood—classic, but ultimately doomed, attempts to reduce modal notions to observable patterns in the actual world. Such patterns are at best good evidence for the corresponding modal facts, and sometimes they are not even that. Against frequentism: consider the fair coin that lands heads every time, or indeed anything other than half of the time, thus ‘acting’ biased. Against the Humean theory: consider effects of a common cause, or spurious correlations, that merely simulate causal relations. Against the naïve regularity theory: consider accidental regularities, such as the fact that every U.S. president since 1821 elected in a year starting in ‘1’ and ending in ‘0’ who was not an actor (a rather imperfect one), died in office. And so it is with the betting interpretation: witness again the flamboyant football team owner, or the puritan. Moreover, the modal notions are meant to *explain* the corresponding patterns; but they can’t do that if they *are* the patterns, a point that Armstrong (1983) has emphasized. And so it is with the betting interpretation: having a certain degree of belief should explain why you bet the way you do, so it can’t just *be* your betting the way you do.

Finally, sometimes the requisite patterns are wholly absent from the actual world, but the underlying facts to be analysed are present nonetheless. This is so obvious in the case of operationalism that *actualist* operationalism was always a complete non-starter: of course there are objects (namely: most of them) that have temperatures but that are never measured for temperature. For some reason, analogous objections have seemed less decisive for the other analyses, even finding some resistance still: the problem of the single-case, or no cases, for frequentism; singular causation for Humeanism;

---

<sup>1</sup>Thanks to Daniel Stoljar for suggesting this last objection. See Block (1981) for further discussion of all three arguments against behaviorism.

uninstantiated laws for the naïve regularity theory (see Tooley 1987). And so it is with the betting interpretation. Indeed, actual bets do not even obey the basic closure conditions of probabilities. You may have placed a bet on a particular horse winning a race, and a bet on a spin of a roulette wheel, but you have never placed one on their *disjunction* or *conjunction*—we are willing to bet.

### Going hypothetical

And so the operationalist/frequentist/Humean/naïve regularity theorist/betting interpretationist is forced to look beyond the actual measurements, actual patterns, actual regularities, actual bets. This introduces a modal element that sits rather uncomfortably with the empiricist/positivist scruples that motivated the positions in the first place, but at least it's an honest acknowledgment that those damned modalities simply won't go away. The operationalist looks to hypothetical measurements; the frequentist looks to virtual sequences of outcomes (Reichenbach 1949, von Mises 1957); the regularity theorist looks to regularities in other possible worlds (Pargetter 1984). And so it goes with the betting interpretation: now identify an agent's credences with the betting prices that the agent *would* make—her betting *dispositions*.

But this helps only up to a point. For if the actual world can be recalcitrant, so can other possible worlds. In neighboring worlds there are still temperatures that go unmeasured (say, at the core of the sun); fair coins that land unfairly; spurious regularities; accidental generalizations. Likewise, there are still propositions on which you may have credences individually, but on which you are not even *disposed* to bet in specific ways, let alone on long disjunctions or conjunctions of them.

Of course, we can idealize: we can look to more remote worlds in which you do have all these dispositions. But to idealize is to fictionalize; there's a risk that we can say whatever we like in the lofty name of 'idealization'. Moreover, it is unclear how the goings-on in such remote worlds could have any bearing on facts about you as you *actually* are. Again, this is a familiar story from the demise of the analyses of the various modal concepts—for example, what bearing does the behavior of some hypothetical infinitely durable coin have on the chance that *this* actual coin lands heads? In some cases we may be able to force you to form the required dispositions—say, by suitable enticement, or by holding a gun to your head. But again, what relevance do these coaxed or coerced dispositions have to your genuine credences? If someone holds a gun to your head and commands you to

form the belief that the moon is made of cheese, it would be rational to do so if you can.

The operationalist definition, in either its actualist or hypothetical guise, thus runs aground. However, while de Finetti claims to give us an operationalist definition, he moves somewhat uneasily between using betting prices to *define* degrees of belief and to *measure* them. Immediately after his insistence on giving such a definition, he continues:

The criterion, the operative part of the definition which enables us to measure it, consists in this case of testing, through the *decisions* of an individual (which are observable), his *opinions* (previsions, probabilities), which are not directly observable. (76)

Testing non-observable opinions in terms of the observable decisions that they give rise to is a way of measuring rather than defining these opinions. This is an admission that there is something *underlying* the decisions, causally related to them but not definable in terms of them. Moreover, de Finetti worries about the relation between utility and money—so will we, shortly—and about agents who care too much or too little about their bets (at very high or very low stakes, respectively). These worries make no sense if credences *are* betting preferences. Very well then; they clearly are *not*. But for all that we have said, betting prices may well *measure* credences.

### Measurement: Jeffrey

Jeffrey's last book (2004) is entitled *Subjective Probability — the Real Thing*. The book opens as follows:

Here is an account of basic probability theory from a thoroughly 'subjective' point of view, according to which probability is a mode of judgment. From this point of view probabilities are "in the mind" the subject's, say YOURS...

How can we get a grasp of this mode of judgment? A natural way is to measure degrees of belief in terms of something observable, like actions, but without reducing them to actions. Note well: *measure*—not *define* or *analyse*. The actions that he has in mind are acts of betting:

if you have an exact judgmental probability for truth of a hypothesis, it corresponds to your idea of the dollar value of a ticket that is worth 1 unit or nothing, depending on whether the hypothesis is true or false. (2)



Note well: “corresponds to”—not “is” or “is defined by”. Betting prices, or at least evaluations of bets, are thus correlated with degrees of belief without being identical to them.

This is more promising than de Finetti’s operationalism. However, the question remains: what *are* degrees of belief? Jeffrey, like de Finetti, considers whether agents will want to accept the relevant bets when the stakes are too high or too low:

The usual way out of this difficulty is to specify that the stakes be small compared with the bettor’s fortune, but not small enough to be boring. The importance of finding a way out is that the ratio of stakes found acceptable is a convenient measure of degree of belief. But we have seen that it is not always a reliable measure. Therefore it seems appropriate to interpret the relationship between odds and utilities in the same way we interpret the relationship between the height of a column of mercury and temperature; the one is a reliable sign of the other within a certain range, but is unreliable outside that range, where we accordingly seek other signs (e.g., alcohol thermometers below and gas thermometers above the range of reliability of mercury). (16–17)

But this only underscores that Jeffrey is not concerned with analyzing degrees of belief here. A reliable sign of something is not the thing itself, and still less is an unreliable sign. In this case: betting prices are *not* the real thing. We now have no definition of the things that we are measuring. The analogy with temperature may not be the best one, since we *do* have a definition of it—a function of mean kinetic energy—and we can use that to give an account of when and why our various ways of measuring it are good ones. No such account of credences is forthcoming. If we now say that there is no possibility of error of our measurements, then we have reverted back to operationalism. But if we admit the possibility and even prevalence of such error, then we risk losing our argument that credences should obey probability theory. Unreliable measurements of them might obey it, but what about *the real things* themselves?

We should admit the possibility of such error—often, and in some cases, large. We saw in the previous section that betting prices provide a poor *analysis* of degrees of belief; now we will see that they may provide a poor *measurement* of them. The football team owner and the puritan suffice to make the point. More generally, placing bets changes the world in various ways, and a rational agent should be sensitive to these changes. For example, you are presumably confident that you will fall asleep tonight; but your

confidence will surely drop if you place a high-stakes bet on your doing so. (Picture the insomniac that you will be at 5 a.m. as you face the imminent prospect of losing your money!) Presumably you have various credences regarding actions that are within your control—as it might be, near certainty that you won't hop on one leg singing 'The Star-Spangled Banner' today; but that will change if you place a bet that rewards you handsomely if you do so. You might even *use* a bet as a way of effecting desirable changes in the world, or in yourself—say, betting with a friend that you will finish your book this year as a way of firming your resolve to do so. Ramsey saw the worry clearly:

the proposal of a bet may inevitably alter [one's] state of opinion; just as we could not always measure electric intensity by actually introducing a charge and seeing what force it was subject to, because the introduction of the charge would change the distribution to be measured. (1990, 35).

The proposal of a bet may alter one's state of opinion in a particularly salient way: it may have interference effects with the proposal of *other* bets. The famous Dutch Book argument for probabilism begins by identifying credences with betting prices (either definitionally, or at least as good measurements). At a crucial step, the *additivity* of such bets is assumed. But if the bets are placed sequentially (as they usually are), then the second bet is placed in the context of a changed world—another bet has been placed—and must be evaluated in that context. It is clearly permissible to revise your betting prices when you know that the world has changed since you initially posted those prices—for instance, when you take the placement of one bet to be correlated with the outcome of a subsequent bet. This is grist for the mill of those who reject the 'package principle' that one should value a package of bets at the sum of their individual prices (see Schick 1986, Earman 1992, Maher 1993).

The non-additivity of bets goes hand-in-hand with the non-linearity of utility in money. Examples are commonplace: you don't have quite enough money to catch the bus home, so you value an even-odds shot at doubling your money more than its expected monetary value. The phenomenon of diminishing marginal utility, or risk aversion, is *so* commonplace that it is often presupposed in economic theorizing, and in interpreting each other. When I say "I bet you a million dollars that Collingwood will win", even odds assumed, I convey to you that my credence in Collingwood's winning is *not*  $1/2$ , but much higher. This is the mirror image of risk-seeking behavior at low stakes.

Of course, de Finetti and Jeffrey were well aware of these phenomena. Their proper resolution comes with utility theory, of which more shortly. But even when credences are well measured by betting prices, there are favorable combinations of bets that are available *only* to probabilistically incoherent agents.<sup>2</sup> Let  $[\$1, X]$  denote the bet ‘\$1 if X, 0 otherwise’, and let ‘Rain’ denote ‘it rains tomorrow’. Consider the following sure-gain books:

- BOOK 1: (1.i) You pay \$0.8 for  $[\$1, \text{Rain}]$   
                   and  
                   (1.ii) you pay \$0.1 for  $[\$1, \neg\text{Rain}]$ .
- BOOK 2: (2.i) You pay \$0.8 for  $[\$1, \neg\text{Rain}]$   
                   and  
                   (2.ii) you pay \$0.1 for  $[\$1, \text{Rain}]$ .

Come what may weather-wise, both books are winners for you: in each case, you pay a total of \$0.9, and receive a guaranteed \$1. Yet only an *incoherent* agent would accept both books. Any coherent agent will find at least one of the \$0.8 bets overpriced and will refuse to buy it, thus refusing the book of which it is a part. But an incoherent agent who assigns, say, probability 0.8 to Rain and 0.8 to  $\neg\text{Rain}$  could happily buy the lot. The incoherent agent does not seem so irrational after all.

The trick here is that these favorable books have, so to speak, a bad part (i) and a very good part (ii), with (ii)’s goodness outweighing (i)’s badness so that overall the package is good. Now, de Finetti and Jeffrey must not say: ‘The rational agent will see the sure gains coming, and so will accept both (i) bets for the sake of the greater good that brings the (ii) bets in its train’. For this is to admit that acceptable betting prices don’t even measure genuine credences. Moreover, if an agent is allowed to see past the status of individual bets for the sake of their overall standing as a package, then that privilege should be bestowed upon the incoherent agent who sees Dutch Books coming and turns them down, while finding acceptable the individual bets of which they are composed. (See Schick 1986.) After all, this is simply to give up the ‘package principle’, just as various opponents of the Dutch Book argument have urged.

Taking stock, we see that there are problems at both stages for the argument for probabilism based on the identification of credences with betting prices. At the first stage, there is a problem with that identification:

---

<sup>2</sup>We thank Adam Elga for this observation, and for a close relative of the example that follows.

credences are not reducible to, nor even necessarily well measured by, betting prices. At the second stage, betting prices that are not probabilities are not obviously irrational.<sup>3</sup>

You may insist that again such problems are resolved with utility theory, and we promise not to keep you waiting much longer for it. But first, we should step back for a moment and scrutinize a foundational assumption that underlies all these theories: the betting interpretation, its fallback position that betting prices merely measure credences, and utility theory. At the core of all of them is the idea that credences should somehow be defined or understood in terms of *preferences*. But credences and preferences are certainly separable in thought, and sometimes in practice. Imagine a Zen Buddhist monk who has credences but no preferences. Gazing peacefully at the scene before him, he believes that Mt. Everest stands at the other side of the valley, that K2 does not, and so on. But don't ask him to bet on these propositions, for he is indifferent among all things. If the monk is conceptually possible, then any account that conceptually ties credences to preferences is refuted. The Stoics thought that one should actually strive for such monk-like indifference. (They, at least, did seem to value *something*—namely truth, and presumably such striving.) Or consider a chronic apathetic who has lost all his desires, but who has kept all his credences. To be sure, these characters are not recognizably like us, although some of us may approximate them over certain domains, and to the extent that we do, bets and preferences more generally ill-reflect our true credences. Not that the ideal Bayesian agent is recognizably like us either; as we have said, to idealize is to fictionalise.

Then there is a problem that seems very relevant to us. Much as beliefs and desires have different *directions of fit* (Smith 1987), so do their next-of-kin, credences and preferences. The goal of a credence is to conform to the way the world is; in the case of a mismatch, a rational agent will typically strive to make appropriate changes in her credence. The goal of a preference is that the world should conform to *it*; in the case of a mismatch, a rational agent will typically strive to make appropriate changes in the world. (Not invariably, in either case—just typically.) Clearly this threatens the prospects for reduction of credence to the preference (and vice versa, for that matter). But plausibly it threatens also the prospects for one measuring the other. As Christensen (2004) emphasizes, credences *represent* the world

---

<sup>3</sup>We do not discuss here the interpretation of the Dutch Book argument that downplays talk of monetary losses, regarding that as merely dramatizing an underlying inconsistency in an incoherent agent's evaluations—see especially Skyrms (1984) and Armendt (1993). That discussion would take us further away from our topic of what degrees of belief *are*.

in a way that may be disconnected from corresponding preferences. We think that some of the counterexamples we have given in this and the previous section are symptomatic of this.

Having thus fired our warning shots, we should move on to the next account of credences that explicitly seeks to ground them in preferences.

## Representation theorems: Ramsey

In various places, Jeffrey goes beyond the measurement view. He refers to Ramsey and his focus on degrees of belief *qua* basis of action, and he cites the Bolker representation theorem that underpins *The Logic of Decision* (Jeffrey 1996). This theorem differs slightly from Ramsey's, but they are essentially alike in laying down axioms on qualitative preference, and then deriving the representability of an agent who obeys these axioms in terms of a probability and a utility function—her preferences represented by maximizing expected utility as calculated by those functions. It is worth noting how little fanfare Jeffrey gives the representation theorem. While the early chapters have much to say about probabilities and utilities, he does not introduce the preference axioms until chapter 9 of *The Logic of Decision*, and even then he gives them rather scant defence. Tellingly, he does not even *state* the representation theorem—one can easily figure out what it must be, but it is nowhere explicitly given in the book. This is hardly surprising—after all, for Jeffrey subjective probability is the real thing, not merely epiphenomenal or a representational artifact. Later on we will side with Jeffrey on this point.

By contrast, for Ramsey the representation theorem takes center stage. It is to his 1926 ground-breaking work that we now turn (Ramsey 1990).

Some of Ramsey's axioms on preferences are reasonable requirements of rationality—indeed, Ramsey goes so far as to call them *consistency* constraints, in keeping with his interpretation of the laws of probability as the *logic* of partial belief. He writes: “Any definite set of degrees of belief which broke them would be inconsistent in the sense that it violated the laws of preference between options. . .” (41). And for the example that he immediately gives, a violation of the axiom of *transitivity*, this is plausible enough. That axiom, however, is especially well chosen to make the point. It is far less clear that a violation of the Archimedean axiom on preferences, say, amounts to *inconsistency*. (Pascal violates it by according infinite utility to salvation—but in doing so he hardly seems to be guilty of a *logical* mistake.) Still less is it clear that the existence of an ethically neutral proposition of probability  $1/2$  is a requirement of consistency, yet this is another of Ramsey's laws of preference.

Be that as it may, with the axioms in place he can prove his famous representation theorem: any agent whose preferences conform to the axioms can be represented as maximizing expected utility as calculated by a utility function  $u$  and a probability function  $p$ . Thus, we have a kind of functionalist account of degrees of belief: they are whatever fills the role of being multiplied with utilities in the expected utility representation.

This account has various virtues over the betting interpretation—notably, it removes the worry about the non-linear utility of money. However, some of the betting interpretation’s problems return. For example, the notion of ethical neutrality is defined in terms of indifference between suitable gambles. But what of the puritan who is indifferent among all gambles, abhorring them all equally? Also, much as the placement of a bet changes the world, and with it, your opinions about the world, so too does the occurrence of a gamble. It is tempting to reply that Ramsey protects himself against this objection by making the outcomes of his gambles entire *worlds*. But far from helping his approach, this move appears to render it incoherent—for what sense is to be made of an *entire world*, complete in all its details, being the result of a gamble?<sup>4</sup> If the gamble is supposed to be part of the ‘world’, then the ‘world’ that is the gamble’s outcome is not really complete after all; but if the world really is complete, then how could it be the outcome of any gamble?

And what of the characterization of credences in terms of their role in the representation theorem? The trouble is that there are *rival* representations, equally faithful to the preference axioms. As Zynda (2000) shows, any agent whose preferences conform to Ramsey’s axioms can be represented as maximizing *non-expected* utility as calculated by a utility function  $u'$  and a non-additive *non-probability* function  $p'$ , by using a non-standard combination rule. Call the numbers output by this non-additive function *schmobabilities*. So the same agent can be represented both in terms of probabilities and in terms of schmobabilities. Which, then, are her degrees of belief? The mere representability of the agent doesn’t settle the matter either way; further argument is required for identifying them with probabilities.

This in turn puts pressure on the associated argument for probabilism. Granted, when the agent’s preferences are right, *something* associated with those preferences obeys the probability calculus. But we need to be convinced that those things are *degrees of belief*. For all we know, degrees of belief are schmobabilities instead.

---

<sup>4</sup>Thanks here to David Chalmers and Brian Weatherson.

More generally, whenever we can compute an expected utility of the form

$$EU(A_i) = \sum_j u(S_j \& A_i)P(S_j)$$

(where the  $A_i$  are actions and the  $S_j$  states), we can recover the same preference ordering with *schemexpected utilities*<sup>5</sup>:

$$SchmEU(A_i) = \sum_j [f(A_i, S_j)u(S_j \& A_i)] \left[ \frac{P(S_j)}{f(A_i, S_j)} \right]$$

Here  $f(A_i, S_j)$  can be any non-zero function you like. The weirdness in the second ‘credence’ bracket is cancelled out by the corresponding weirdness in the first ‘utility’ bracket. But to use the word ‘weirdness’ here is already to prejudge the issue in favor of probabilism. Again, given that we can represent the agent both in terms of quantities that obey probability theory and in terms of quantities that don’t, which are the degrees of belief? Indeed, why should we think that *any* of the representations—probabilistic and non-probabilistic alike—codify the agent’s degrees of belief? Correspondingly, why should we think that degrees of belief are probabilities? Clearly more needs to be said about degrees of belief.

Ramsey does say more. He writes: “a degree of belief is a *causal* property of it, which we can express vaguely as the extent to which we are prepared to act on it. This is a generalization of the well-known view, that the differential of belief lies in its *causal* efficacy. . .” (1990, 170, our emphases). This is music to the ears of functionalists; in fact, Ramsey deserves more credit than he gets in the philosophy of mind literature for being a true pioneer of functionalism. But stare as long as you like at Ramsey’s representation theorem, and you will not find any representation of *causation* in it.<sup>6</sup> So two agents who are identical regarding their inputs into the theorem, their preferences, will be identical regarding their outputs, the utility/probability functions attributed to them, irrespective of what causal relations may be instantiated. Any agent for whom the causal relations stand one way can be simulated by one in which they stand another way, or in which they are absent altogether. Blockhead<sup>7</sup> is as susceptible to a Ramseyian representation as you are—in fact more so, for we may suppose that Blockhead’s

---

<sup>5</sup>We thank Jim Joyce for pointing out this trick.

<sup>6</sup>We are indebted here to Daniel Nolan.

<sup>7</sup>Blockhead was introduced in Ned Block’s (1981) attack on functionalism, although not under that name; the name has become popular since. Block imagines a computer that is able to pass the Turing test despite not being intelligent: it is pre-programmed with sensible responses to all possible conversational prompts.

preferences meet Ramsey's stringent demands, whereas we can be sure that yours do not.

### Radical interpretation: Lewis

Lewis's (1974) important project begins as one of *interpreting* an agent whose language, beliefs and desires are unknown to us, on the basis of the facts about her as a physical system. At the outset, then, it is apparently not about what beliefs, or degrees of belief, an agent actually has, but about how we come to interpret her as having particular beliefs, or degrees of belief. Yet Lewis insists: "I am not really asking how *we* could determine these facts [about the agent's language, beliefs, and desires]. Rather: how do *the facts* determine these facts" (110, emphasis in the original). Well, let's see.

He argues that we should ascribe those degrees of belief to an agent—Lewis calls him "Karl"—that fit maximally well with a number of principles. Two of them are relevant here: the *principle of rationalization* and the *principle of charity*. The principle of rationalization says that an agent should be ascribed those degrees of belief that would rationalize his behavior:

Karl should be represented as a rational agent; the beliefs and desires ascribed to him... should be such as to provide good reasons for his behavior..." (337)

Lewis explains that this means that we should ascribe those beliefs and desires that would make the agent's actions maximize expected utility:

Take a suitable set of mutually exclusive and jointly exhaustive propositions about Karl's behavior at any given time; of these alternatives, the one that comes true... should be the one (or: one of the ones) with maximum expected utility according to the total system of beliefs and desires ascribed to Karl at that time... (337)

Presumably when ascribing beliefs and desires to Karl on the basis of his behavior, we should be careful not to *hyperrationalize* that behavior. Some of it is not intentional action; sometimes he just idly drums his fingers unconsciously, and we should not read too much into that. More disturbingly, the principle of rationalization presupposes that we already know what rationality consists in (i.e. maximizing expected utility). This makes the principle unfit to figure in an *argument* for probabilism, a thesis about what rational degrees of belief are. We will return to this point at the end of this section.

The principle of charity says that an agent should be ascribed those degrees of belief that he ought to have. Lewis quotes Davidson, who contends



that when we interpret an agent we should “find him consistent, a believer of truths, and a lover of the good” (Davidson 1970, 3). Davidson really means: a believer of what *we* believe, a lover of what *we* love. After all, we are not omniscient, so what we think is true and good might not actually be so. So when ascribing beliefs and values to others, we must be guided by our own opinions about the true and the good. However, Davidson’s principle needs some finessing, Lewis argues. The mark of a rational agent isn’t that he only believes true propositions, but that he believes what is reasonable given his available information. The agent, however rational, might not have reached our own elevated state of insight and wisdom if he hasn’t had all the evidence that we have. So his lack of evidence should be taken into account, Lewis thinks: we should ascribe those beliefs and desires that we *would* have, were we in his place.

Karl should be represented as believing what he ought to believe, and desiring what he ought to desire. And what is that? In our opinion, he ought to believe what *we* believe, or perhaps what we would have believed in his place; and he ought to desire what we desire, or perhaps what we would have desired in his place. (336)

One wonders how the interpretation is supposed to work if it is Karl who has the more elevated state of insight and wisdom, since we haven’t had all the evidence that he has; but let that pass. In any case, the upshot is: we should ascribe to Karl what it is rational for him to believe by our lights, given the evidence that he has. This is as true of degrees of belief as it is of beliefs.

At this point it appears that Lewis *is* “really asking how *we* could determine these facts [about the agent’s language, beliefs, and desires]”. The principle of charity ineliminably involves *us*. And who is that, we might ask? Bob, Carol, Ted, and Alice all believe different things—their stocks of evidence are different, their inductive methods are different, and some of their beliefs may have little basis in evidence or method at all (as it might be, a basis in early upbringing instead). They all desire different things, both intrinsically and instrumentally. They each set to work interpreting Karl, and come up with four different representations of him, and in particular attribute four different credence functions to him. What, then, is his true credence function? Certainly not all of them; and chances are that it is none of them. Will the real Karl please stand up! It will hardly do to say that to the extent that the interpretations differ, there is indeterminacy in Karl’s mental state. It’s not as if the more friends Karl has interpreting him, and hence the more disagreement in their interpretations, the less determinacy there is in what he believes and desires. Who would want friends?

And when we ascribe to Karl what it is rational for him to believe by our *lights*, what are they? Our own *credences*, presumably. But ‘credences’ were the very things we sought to analyze, so we do not have a reductive analysis yet. One wonders, moreover, how the analysis is to apply to ourselves. It’s not as though Karl has to reciprocate the favor and interpret us, or if not Karl then someone else, in order for us to have credences. Interpretation just does not seem to be the point. If we say that by definition a person has those degrees of belief and utilities that facilitate the best interpretation of him, we confuse epistemology with metaphysics. An interpretation according to Lewis’ principles might be our best hypothesis regarding the agent’s credences, but that doesn’t ipso facto guarantee the truth of that hypothesis. But Lewis claims that the best interpretation of Karl gives us his actual degrees of belief—indeed, Lewis even suggests that this is *analytically true*. But this is too good to be true, let alone analytically so. Compare an astronomer who claims that it is *analytically true* that if our best theory says that planets follow elliptical orbits, then that’s exactly what they do. That’s not good science; that’s hubris.

What we want, then, is the *best possible* interpretation of Karl—God’s interpretation, we might say metaphorically. This solves the problem of disagreeing interpreters—we should all defer to God’s interpretation. It also solves the problem of the truth that the interpretation is correct being analytic—God is entitled to some hubris. But then *interpretation* falls by the wayside. We might as well say simply that God knows directly what Karl’s credences are: they are part of the ⟨probability, utility⟩ profile that *in fact* best rationalizes his behavior, renders his credences reasonable, and so on. Then mental states are no more a matter of interpretation than planetary orbits are.

Moreover, even God may be frustrated in his attempts to interpret us. Without the metaphor: the best interpretation available may not be much good. As the saying goes, nobody is perfect. But our imperfections are no impediment to our having credences. Indeed, it is partly *because* of our imperfections that we have intermediate credences on logical matters, where probabilism shuns them. In any case, our credences are often unreasonable, flummoxing a charitable interpreter, but so be it.

Or there might be multiple conflicting interpretations of us that are equally good. The beliefs that would best rationalize our actions are often false; our behaviors can often be *rationalized* only by an interpretation of our beliefs that is *uncharitable*. And often our beliefs can be interpreted charitably, but our behaviors not rationalized. This then forces a trade-off between the principles of rationalization and charity. Consider an example

due to Maher (1993), in which he imagines himself betting on successive tosses of a coin as if in the grip of the gambler's fallacy; but he would also bet that there are no causal relations between these tosses. He's right: there are none, so the principle of charity would bid us to attribute to him a probability function according to which the tosses are independent. But how are we to square that with his gambler's-fallacious preferences? To rationalize those, we have to retract our charitable attribution. If any interpretation reaches a reasonable threshold of adequacy, then many do. Lewis ends his paper by doubting that there could be multiple perfect interpretations of a single agent. Grant that to him; but there could be multiple *imperfect* interpretations of an imperfect agent, the equal best of a bad lot.

These points generalize: surely we can have degrees of belief, degrees of desire, and preferences even when we defy a single perfect interpretation. Do we then have these mental states to a lesser extent than does an agent who may be attributed a perfect interpretation? At an extreme, does an agent who foils our attempts at interpretation lack a mental life altogether? Think again of our friend the Zen Buddhist monk staring silently at Mt. Everest. We could interpret him as being at one with the universe, or as having no inner life whatsoever. Nothing in his behavior settles the issue, and Lewis's principles get no traction. *God knows what his true mental state is!* More realistically, there are all too many behaviors that we should *not* rationalize and interpret charitably—just watch the nightly news if you need examples.

This problem regarding the account of degrees of belief (what we have called “stage i”) in our introduction) quickly becomes a problem regarding any argument for probabilism based on it (“stage ii”). It would trivialize probabilism if it were *impossible* for credences to violate the probability calculus. The concern is that an attribution of such credences will, according to the probabilist, run afoul of both the principles of charity and of rationalization, and so will never be admissible. The interpretivist presumably thinks that it would be more charitable to attribute some credence profile that obeys the probability calculus; and the notion of expected utility that underlies any rationalization assumes probabilities as inputs, not schmobabilities that violate probability theory. There are agents whose credences violate the probability calculus (Exhibit A: you; Exhibits B and C: us). But interpretivist accounts seem unable to say this: either they interpret such agents as having credences that are automatically probabilities (else the principles for interpretation are not applied), or they interpret them as not having credences at all.

Furthermore, the application of the principles of charity and rationalization presupposes that we already know what a charitable and rationalizing

interpretation is. The interpretivist presupposes that the appropriate normative standard is probabilism. A schinterpretivist presumably thinks that it would be more charitable to attribute some credence profile that obeys the *schmobability* calculus; and the notion of expected utility that underlies any rationalization assumes *schmobabilities* as inputs, not probabilities that violate schmobability theory. This leaves us with a stalemate between probabilism and schmobabilism until we have an independent argument that it is the interpretivist who applies the correct principles for interpretation.

We take it that Lewis's talk of 'interpretation' is really just a heuristic. To be sure, it is a more sophisticated heuristic than, say, the betting interpretation. But still it leaves the full story about credences, and with it a corresponding story about why they should be probabilities, untold.

### Interpretivism: Maher

Maher is another interpretivist. While his overall position is much in the spirit of Lewis's, Maher puts more emphasis on the decision theoretic representation theorem, and on the associated argument for probabilism. He also argues that rationality requires more of an agent than merely having preferences that can be given an expected utility representation, and that there is a standard of rationality for degrees of belief more fundamental than one based on expected utility maximization:

I think we should allow that probabilities and utilities can themselves be irrational; and when they are, it will not generally be rational to maximize expected utility relative to them. Thus I view the principle of maximizing expected utility as elliptical; to get a literally correct principle, we must add to the principle of expected utility the proviso that the probabilities and utilities are themselves rational. (29)

Bayesianism must thus be *qualified*, Maher argues, so that degrees of belief and utilities are rational in themselves and not merely in how they combine with each other. Return to Maher's gambler's fallacy example, now with the little extra detail that he gives it. Imagine that most of his betting preferences succumb to the gambler's fallacy. But he also has a single further preference that does not fit: he would bet that there is no causal relation between one toss and the other.<sup>8</sup> Which preference(s) should he give up?

---

<sup>8</sup>Actually, this way of speaking is somewhat infelicitous, although it is faithful to Maher's text. We have just described a *betting disposition* rather than a *preference*. A disposition may be explained by a corresponding preference, but they should not be identified.

Clearly those that are based on the gambler's fallacy, argues Maher, and we agree. And yet those are in the majority, and his reasonable preference is in the minority. Maher surmises: "Unqualified Bayesianism here endorses majority rule: It says that most of my preferences are rational and the few that diverge are irrational. But when I reflect on the conflict, I might well decide that it is the minority that are right and the majority that are wrong" (32).

Maher is assuming here that the problem of *attributing* a probability and utility function has already been resolved, also by majority rule. A probability/utility function pair has been fitted according to which the gambler's fallacious preferences maximize expected utility. It is unclear to us why the attribution problem must be resolved that way. Certainly, nothing in the representation theorem approach settles it. The theorem gives us a conditional in one direction: *If* all of his preferences obey certain axioms, *then* he is interpretable as maximizing expected utility relative to some  $u$  and  $p$ . But if, as must be the case in Maher's example, *not* all of his preferences obey the axioms, then all bets are off. (If they *all* obeyed the axioms, then Bayesianism's endorsement would be whole-hearted: it would say that *all* of his preferences are rational!) At a first pass, then, the representation-theoretic resources for attribution would seem to go silent here.

Silence is surely intolerable; after all, in the example he does not have the mental life of a rock! We think it would be more natural, and still in the spirit of the representation theorem approach, to find maximal subsets of his preferences that obey the axioms, and to tailor probability/utility functions piece-wise to those. To be sure, this will yield a kind of indeterminacy in the probability/utility attribution to him in the example—the best that can be said is that he has one such pair over *this* domain of propositions, and another such pair over *that* domain, and no single pair over both. But better indeterminacy than silence.

Maher would apparently not find this satisfactory either, given his endorsement of the 'majority rules' approach to probability attribution (at least in this example). He then wants to judge some of the preferences, so interpreted, as irrational (the majority, in the example), and others as rational (the minority, in the example). Presumably the minority is rational because it maximizes expected utility relative to *rational* credences: those that reflect no causal influence between tosses are rational, while those that

---

And speaking here of *a* betting disposition, in the singular, is also infelicitous. Presumably he has a raft of such dispositions: being disposed to bet on there being no such causal connection *at such-and-such stakes and at so-and-so prices*, for potentially infinitely many instances of the stakes and prices.

reflect the gambler's fallacy are irrational. This apparently presupposes that the attribution of credences and utilities should be done piece-wise as we suggested. If the agent has only one pair of probability and utility functions—those that rationalize the majority preferences—he would simply *not have any* credences rationalizing the minority preferences.

In any case, even the credences that Maher rightly judges to be irrational are still *probabilities*. But what of credences that violate the probability calculus? He seems to lack the resources for attributing such credences—after all, they certainly are not delivered by the representation theorem. He then seems to be in the odd position that he *can* attribute crazy credences (say, high credences to alien abductions) provided that they are probabilities, but he cannot attribute sensible credences that are not probabilities (say, a credence of 0.8 that it will rain tomorrow, and 0.15 that it won't). And it seems even more odd that Maher would insist that the latter attribution is impossible while the former one may be correct, given his emphasis on rationality as involving a standard of reasonableness that goes beyond probabilistic coherence.

## Primitivism<sup>9</sup>

Despite our criticisms, we are hardly in a position to claim any intellectual or moral high ground compared to the various authors that we have surveyed. For at least they have all attempted to tell us what degrees of belief are, while we have not and will not here.

To be sure, they have succeeded in providing much illumination of 'degree of belief'; but we have argued that *qua* analyses, their attempts have failed. One might take a hint from this failure: that the project of analysing degrees of belief was misguided from the start. One might run a Laudanian pessimistic meta-induction, based on this failure, predicting that all such analyses *will* fail. One might even have a Quinean suspicion that the whole project of conceptual analysis is misguided, and conclude that all such analyses *must* fail. Skeptics of both stripes have already drawn such conclusions regarding other concepts that have proved resistant to analysis. If 'degree of belief' turns out to be elusive to the conceptual analyst, at least arguably it's in good company.

But let us not rest content merely with *via negativa* arguments for primitivism about 'degree of belief'. For perhaps we should have seen this parti-

---

<sup>9</sup>This section has profited considerably from discussions with Alex Byrne, Frank Jackson, Carrie Jenkins, Peter Menzies, and Daniel Nolan.

cular failure coming. After all, it should come as no surprise if ‘degree of belief’ is *especially* resistant to analysis, since it is such a basic epistemic notion. Said more positively, perhaps it makes an especially natural primitive. Like Gertrude Stein’s rose, a credence is a credence is a credence—and not something else.

It is a strangely neglected topic in analytic philosophy—an enterprise up to its neck in attempted conceptual analyses—what make good primitives for such analyses in the first place. Let us say something about this book-length topic in a page or two. First let us distinguish *ontological* from *conceptual* primitives. The former involve the fundamental entities and properties in the universe (as it might be, electrons and charge, although even they may not be fundamental enough); the latter involve our fundamental concepts, the most basic building blocks in our thinking. We turn to science to inform us about the former, to philosophy to inform us about the latter. Our question, then, is what makes a particular choice of a conceptual primitive (or a set of such primitives) propitious.

Here is a wish list of criteria—not jointly sufficient, and none necessary—for a good choice of primitives. For starters, it should be well understood (at least compared to its rivals). It should be natural, rather than gerrymandered, carving the domain in question at its joints, classifying like entities with like. It should cut possibility space finely, compared to its rivals, allowing more distinctions that we care about to be made. As a result, it will offer more opportunities for reduction than a poor choice of primitives. But it should not cut unnecessarily finely, carving at the joints and then some; otherwise conjunctions of good primitives would automatically make better primitives. It should be responsive to our conceptual needs but not gratuitously outrun those needs.

Above all, primitives should be judged by the total theory in which they appear. They should foster systematisation, figuring in economical but general theorizing of the domain in question. This evokes the *best system* account of the *laws of nature*, associated with Mill, Ramsey and Lewis, as theorems of the true theory of the universe that best balances simplicity and strength. A theory’s primitives enjoy a reflected glory from whatever virtues the theory may have. A good theory, for example, will be qualitatively parsimonious (Lewis 1973), not trafficking in more primitives than it needs to. The best theory that we can have is the best guidance we can get to our choice of primitives. It is by its contribution to a virtuous total theory that a primitive earns its keep. This also resonates with the view of explanation as involving *unification* (Friedman 1974, Kitcher 1989). More specifically,

well-chosen primitives should figure in explanations and predictions of other phenomena, rather than the other way round.

This doesn't mean that a given choice of primitives has to be unique, exclusive of other choices. Euclidean geometry can be formulated with points as primitive, or with lines as primitive. The logical connectives can be interdefined, with '¬' and '∨' as primitives, or '¬' and '&', or other combinations; or they can all be defined in terms of the Sheffer stroke ('nand'); or in terms of its dual, the Sheffer dagger ('nor'). Similarly, the interdefinability of the quantifiers and of the modal operators shows that we have some freedom over which we take as primitive. Said in Mill/Ramsey/Lewisian terms, there may be multiple equally good systematisations of a given domain. But a well-chosen primitive should appear in some such systematisation.

This is the job description; we submit that 'degree of belief' is the ideal candidate for the job of being an epistemological primitive.

In our introduction we compared probabilism with traditional epistemology, and its categories of 'knowledge', 'justification' and (binary) 'belief'. While attempting to reduce 'knowledge' to other notions is one of the biggest industries that philosophy has ever seen, and while 'justification' has also been regarded as needing reduction, 'belief' has been more readily spared these reductionist aspirations. Apparently it has been regarded as fit to be a primitive concept, at least in epistemology. 'Degrees of belief' have, we think, equal claim to being regarded as natural. But they are, if anything, even more fundamental. Clearly, they allow finer-grained distinctions to be made among epistemic states: they may take any of the uncountably many real numbers in  $[0, 1]$  as values, and so are more discriminating, to put it mildly, than binary beliefs (which we could represent simply with 0 or 1). But they are not gratuitously discriminating: every distinction among degrees of belief could in principle make a difference to some claim of confirmation or some decision. And once 'degrees of belief' are in our conceptual storehouse, there is some hope that 'beliefs' may be reduced to them—e.g. 'to believe that  $p$  is to assign degree of belief to  $p$  above some contextually salient threshold', a version of an idea that Bovens and Hawthorne (1999) attribute to Locke. But it is hard to see how the reduction could go the other way—there is simply too much structure in 'degree of belief' that is missing from ungraded 'belief'.

And placing the structure into the contents of belief would be to *misplace* it. A credence in  $X$  is not to be analyzed as a full belief regarding the probability of  $X$ , in any sense of 'probability'. You may give credence  $1/2$  to it raining tomorrow, without believing that the probability of rain is  $1/2$  in any sense—e.g. you may believe that the objective probability is 0 or 1, although



you don't know which; you may lack introspective access to your own subjective probability; you may lack the concept of probability altogether; you may have the concept but think that it never applies; and so on. Moreover, binary 'belief' is far too coarse-grained to offer reductions of various other related notions, whereas 'degree of belief' apparently fits the bill admirably; we will discuss the notions of 'confirmation' and 'rational decision' shortly. So the orthodoxy in epistemology, to the extent that there is one, is a rather odd one: supposedly, 'degree of belief' is in need of reduction, but 'belief' is not. We see matters exactly the other way round.

Nor can 'degree of belief' be reduced to the other degreed notion that is the stock-in-trade of the Bayesian, desirability, or more fundamentally, to preference. The alleged connections between credences and attitudes towards betting, and preferences more generally, are tenuous (recall how measuring one's attitudes can change them), defeasible (recall the football team manager), and circumscribed (recall the problems of low and high stakes). It isn't just that we claim to have cast doubt on some promising contenders for such reduction in our critiques above (although this is also true). More generally, recall that we had a principled skepticism for the prospects of any such reduction, based on the opposite directions of fit of degrees of belief and degrees of desire, and the conceptual possibility of agents who could have the former without the latter (the chronic apathetic, the monk). So preference-based accounts of degrees of belief are unfit to serve as analyses of them, and thus unable to bear the burden of the arguments for probabilism that have been placed on them.

Much as preference-based accounts of degree of belief get the direction of fit wrong, we believe that they also get the order of explanation wrong (although we realize that this is controversial). Why do you prefer the status quo to paying 50 cents for a dollar bet on the Republicans winning the next U.S. election? Because your credence that the Republicans win is less than  $1/2$ . Why do you prefer buying fire insurance for your house rather than leaving it uninsured? Because you have some small-but-non-negligible credence that your house will burn down, some large disutility attached to it doing so, and some smaller disutility to paying the insurance premium, such that the expectation calculations favor your getting insurance. Or recast these examples in terms of predictions: Knowing your credence for the Republicans winning, we can predict that you will turn down that 50-cent bet when offered it. Knowing your credence and disutility for your house burning down, and your disutility for paying the premium, we can predict that you will buy the insurance. The examples generalize. The order of explanation, or of prediction, is not: 'These are your betting dispositions,

or preferences; thus, those are your credences and desirabilities.’ Rather, it is: ‘These are your credences and desirabilities; thus, those are your betting dispositions, or preferences.’ Recall how the first eight chapters of Jeffrey’s classic book on decision theory takes credences and desirabilities for granted; the preference axioms only appear later in the piece, and the associated representation theorem is left tacit.

Moreover, credences explain far more than just preferences, as Christensen (2004) forcefully argues. They explain other aspects of behavior besides those that are preference-related. Consider a golfer facing a crucial putt, whose high confidence of sinking it steadies his hands, thereby making it more likely that he will indeed sink it. And credences explain much more than behavior. Consider their role in causing other psychological states, such as happiness (e.g. when one is confident of success), regret (e.g. when one has high credence in certain counterfactuals about what would have transpired had one acted differently), embarrassment (e.g. when one comes to realize that it was the dean’s car that one was tap-dancing on at last night’s drunken party), and so on. Or consider their role in explaining psychological processes, such as inference, as Christensen observes. Preference-based accounts of credence fixate on just one kind of role that credences play, when in fact the roles are many and varied.

So our proposal is to take Jeffrey’s dictum that subjective probability is ‘the real thing’ even more seriously than Jeffrey himself may have done, and to use the unanalyzed notion unapologetically and without compunction in our theorizing and explanatory enterprises. Science is hardly stymied by its failure to provide analyses of its most basic concepts—charge, time, mass, spin, distance, and so on. And the social sciences progress well enough without adequate analyses of their concepts—agent, revolution, nation, war, memory, regret, and what have you—even though we are confident that these things are not to be found in the fundamental inventory of the universe. Closer to home, arguably the majority of our most interesting philosophical concepts are unanalysable<sup>10</sup>; indeed, their unanalyzability may be *part* of what makes them philosophically interesting. Still closer to home, some philosophers insist on the unanalyzability of certain specific concepts kindred to ‘degree of belief’—think of Williamson on ‘knowledge’ (a kindred epistemic concept), or Sober on ‘objective probability’ (a kindred probabilistic concept).<sup>11</sup> If our concept of ‘degree of belief’ continues to be used without being successfully reduced, once again it’s arguably in good company.

---

<sup>10</sup>There are shades of this idea in the writings of Jerry Fodor, Mark Johnston and others.

<sup>11</sup>See Williamson (2000); Sober (2005).

Furthermore, taking ‘degrees of belief’ to be primitive does not preclude us from saying many informative things about them. Indeed, some arguments for probabilism apparently *take* them as primitive, and then appeal to some norm that we putatively know about them, perhaps tacitly. For example, calibrationists argue that credences should strive to track corresponding relative frequencies (e.g. van Fraassen 1984); Joyce (1998) argues that they should strive to track corresponding truth values; Hájek (MS) argues that they should strive to track corresponding chances. These norms then figure directly as premises in arguments for probabilism, without any thought of reducing credences to something else. Indeed, the remarkable success of Bayesianism indirectly provides a pragmatic argument for probabilism: that very success is a reason to accept it. And we have argued that credences have a particular direction of fit, and we have used our tacit knowledge about them to support various counterexamples throughout this paper to the analyses on offer. Primitive they may be, but that does not render them obscure.

‘Degree of belief’ is also an intuitive notion in folk psychology, familiar from everyday speech and thought. In this respect it’s nothing like ‘spin’, and not much like ‘charge’ either. Still less is it like avowedly introduced technical notions, such as ‘Banach space’ or ‘Turing machine’. It may *look* technical by the time probability and decision theorists have got their hands on it and massaged it. But they have only formalized and regimented a perfectly mundane concept. After all, we have various ways in English (and, we hazard to guess, every natural language) for conveying our degrees of belief. Think of the spectrum of phrases that we have at our disposal: “I’m certain that  $p$ ”, “I’m almost certain that  $p$ ”, “I’m extremely confident that  $p$ ”, “I’m moderately confident that  $p$ ”, “I’m fairly confident that  $p$ ”, ... all the way down to their duals at the other end: “I’m certain that not- $p$ ”, and so on. To be sure, that gives us at best an ordering—presumably only a partial ordering—of propositions according to the confidence that we accord them. That’s where the regimentation comes in. We fix a zero point at maximal uncertainty, a unit at maximal certainty, and get a handle on the intermediate values thanks to the additivity axiom. But this only adds structure and nuance to a concept of ‘degree of belief’ that we already had. We display our mastery of that concept by:

- applying it appropriately—e.g. in expressing our own opinion;
- performing inferences properly using it—e.g. in inferring that not- $p$  is probable if  $p$  is improbable;

- knowing what counts as evidence for attributions involving it—e.g. linguistic utterances, and yes, often but not invariably betting behavior. And we also recognize the limits of such evidence, knowing when it can be trumped; many of our arguments against the betting interpretation deployed such knowledge, and all of them would surely be intelligible to reflective folk. Thanks to our antecedent understanding of the concept, we can intuit when betting prices are good measures of degrees of belief, and when not.

This talk of the folk suggests that degrees of belief can be understood along the lines of the so-called *Canberra plan*. “Ramsify over the folk platitudes regarding the concept” is the credo where we are based. (The classic manifesto is Jackson 1997.) We might look to the best philosophical work on degrees of belief that comes to mind—and we’ve seen a fair bit of it—when trying to provide the Ramsey sentence for ‘degree of belief’. “It is the thing that should obey the probability calculus (and thus provides an interpretation of it), that is often but not always measured by betting dispositions, that guides decision, that underpins relations of confirmation between evidence and hypotheses, that is often revealed in verbal reports, and so on.” So it may be, although Russell’s famous quip also comes to mind: “The method of ‘postulating’ what we want has many advantages; they are the same as the advantages of theft over honest toil.” (Russell 1919: 71) Say, if you like, that to Canberra-plan ‘credence’ is to offer an analysis of it; it is certainly not the traditional project of reducing ‘credence’ to something else. We prefer to say that it is to take ‘credence’ as primitive, and to list a few things that we know about it (and while we’re at it, why stop with platitudinous things?). Merely Ramsifying over them, and doing no more, strikes us as a way of endorsing primitivism until—fingers crossed!—someone does the hard work of telling us what plays the credence role. And if we’re right, so much the better for primitivism.

Further defense of primitivism about credences, in the spirit of Canberra planning, comes from the fact that, despite the failure of the various analyses of credences, probabilism is a flourishing research program. We may not be able to say what a degree of belief *is*, but that doesn’t prevent us from saying many substantive things about its role in our conceptual economy. Compare again charge: we recognize it by the various equations of physics in which it appears, even if we can’t reduce it to something else. If you need any convincing of just how much work can be done with a primitive notion of ‘degree of belief’, go look at Bayesian tomes such as Earman (1992) and Howson and Urbach (2006). To be sure, they gesture at some of the analyses

that we have rejected, the betting interpretation in particular. But notice how many of the fruits of Bayesianism can be reached without any appeal to these analyses. Bayesianism illuminates key episodes in scientific theorizing; it helps explain why diverse evidence is a good thing; it captures what was right in the HD account of confirmation, while also diagnosing what was wrong; it goes a long way to resolving such old chestnuts as the grue paradox, the raven paradox, and the Quine-Duhem problem; and much more.

And so we come back full circle to our major reason for being happy to regard ‘degrees of belief’ as primitive: they are part of our best systematisation of epistemology. Indeed, we are tempted to go so far as to say that probabilism codifies the laws of (normative) epistemology. This has the air of metaphor, but we think that there is a good measure of literal truth to it. A remarkably simple theory—in essence, three axioms that you can teach a child—achieves tremendous strength in unifying our epistemological intuitions. Rather than cobbling together a series of local theories tailored for a series of local problems—say, one for the grue paradox, one for the raven paradox, and so on—a single theory in one fell swoop addresses them all. While we’re at it, the same theory also undergirds our best account of rational decision-making. These very successes, in turn, provide us with an argument for probabilism: our best theory of rational credences says that they obey the probability calculus, and that is a reason to think that they do.

None of these successes awaits some further analysis of ‘degree of belief’. Indeed, it would seem rather odd to be told that some result in confirmation theory is essentially one about betting behavior, or one about expected utility maximization. It would be as if the entire history of science were some protracted exercise in reasoning along the lines of Pascal’s Wager! Give ‘degrees of belief’ their due: they do not need to be subordinated to something else.

## Conclusion

So at the end of the day, is our position as badly off as those that we criticize? After all, we refuse to say what a degree of belief is. However, we don’t *try* to say what a degree of belief is. We don’t try to, because as we have argued, we don’t need to.

This is not to say that we should give up trying to analyze ‘degree of belief’. We can only applaud Ramsey, de Finetti, and co for their efforts. If we did not welcome someone coming up one day with a satisfactory analysis, we should turn in our badges as philosophers. But in the meantime, we can get by well enough without such an analysis. After all, we’ve done so up to now.

**Acknowledgements.** We thank John Bigelow, Alex Byrne, Kenny Easwaran, Adam Elga, Chris Hitchcock, Franz Huber, Frank Jackson, George Kahramanis, Patrick Maher, Ralph Miles, and especially Peter Godfrey-Smith, Carrie Jenkins, Stephan Leuenberger, and Daniel Nolan, for very helpful comments and discussions.

## References

- [ARMENDT, BRAD (1993)] ‘Dutch Books, Additivity, and Utility Theory’, *Philosophical Topics* 21:1–20.
- [ARMSTRONG, D. M. (1983)] *What is a Law of Nature?*, Cambridge, Cambridge University Press.
- [BLOCK, NED (1981)] ‘Psychologism and Behaviorism’, *Philosophical Review* 90:5–43.
- [BOVENS, LUC, and JAMES HAWTHORNE (1999)] ‘The Preface, The Lottery, and the Logic of Belief’, *Mind* 108:241–264.
- [CHISHOLM, RODERICK (1957)] *Perceiving*, Ithaca, New York, Cornell University Press.
- [CHRISTENSEN, DAVID (2004)] *Putting Logic In Its Place*, Oxford, Oxford University Press.
- [DAVIDSON, DONALD (1970)] ‘Mental Events’, in L. Foster and J.M. Swanson (eds.), *Experience and Theory*, Cambridge, MA, University of Massachusetts Press.
- [DE FINETTI, BRUNO (1990)] *Theory of Probability*, Vol. I (first edition 1974), New York, Wiley.
- [EARMAN, JOHN (1992)] *Bayes or Bust?*, Cambridge, MA, MIT Press.
- [FRIEDMAN, MICHAEL (1974)] ‘Explanation and Scientific Understanding’, *The Journal of Philosophy* 71:5–19.
- [HÁJEK, ALAN (MS)] *A Puzzle About Degree of Belief*.
- [HOWSON, COLIN, and PETER URBACH (2006)] *Scientific Reasoning: the Bayesian Approach* (first edition 1989), La Salle, Ill, Open Court.
- [JACKSON, FRANK (1997)] *From Metaphysics to Ethics*, Oxford, Oxford University Press.
- [JEFFREY, RICHARD (1966)] *The Logic of Decision*, Chicago, University of Chicago Press; 2nd ed. 1983.
- [JEFFREY, RICHARD (2004)] *Subjective Probability—The Real Thing*, Cambridge, Cambridge University Press.
- [JOYCE, JAMES M. (1998)] ‘A Non-Pragmatic Vindication of Probabilism’, *Philosophy of Science* 65:575–603.
- [KITCHER, PHILIP (1989)] ‘Explanatory Unification and the Causal Structure of the World’, *Minnesota Studies in the Philosophy of Science*, 410–505.
- [LEWIS, DAVID (1973)] *Counterfactuals*, Cambridge, MA, Harvard University Press.
- [LEWIS, DAVID (1980)] ‘A Subjectivist’s Guide to Objective Chance’, in *Studies in Inductive Logic and Probability*, Volume II, Berkeley and Los Angeles, University of California Press, 263–293; reprinted in *Philosophical Papers*, Volume II, Oxford, Oxford University Press, 1986, 83–132.
- [LEWIS, DAVID (1974)] ‘Radical Interpretation’, *Synthese* 27:331–44; reprinted in *Philosophical Papers*, Vol 1, New York and Oxford, Oxford University Press 1983.
- [MAHER, PATRICK (1993)] *Betting on Theories*, Cambridge, Cambridge University Press.

- [PARGETTER, R. (1984)] 'Laws and Modal Realism', *Philosophical Studies* 46:335–47.
- [PUTNAM, HILARY (1965)] 'Brains and Behavior', in R. J. Butler (ed.), *Analytical Philosophy*, Second Series, New York, Barnes & Noble 1–19.
- [RAMSEY, F. P. (1990)] 'Truth and Probability', in D.H. Mellor (ed.), *F. P. Ramsey: Philosophical Papers*, Cambridge, University Press, Cambridge (first published in R. B. Braithwaite (ed.), *Foundations of Mathematics and other Essays*, London, Routledge & P. Kegan, 1931, 156–198; reprinted in H. E. Kyburg, Jr. and H. E. Smokler (eds.), *Studies in Subjective Probability*, 2nd ed., Huntington, NY, R. E. Krieger Publishing Company, 1980, pp. 23–52).
- [REICHENBACH, HANS (1949)] *The Theory of Probability*, Berkeley, University of California Press.
- [RUSSELL, BERTRAND (1919)] *Introduction to Mathematical Philosophy*, London, Unwin.
- [SCHICK, FREDERIC (1986)] 'Dutch Bookies and Money Pumps', *The Journal of Philosophy* 83:112–119.
- [SKYRMS, BRIAN (1980)] *Causal Necessity*, New Haven and London, Yale University Press.
- [SKYRMS, BRIAN (1984)] *Pragmatics and Empiricism*, New Haven, Yale University Press.
- [SMITH, MICHAEL (1987)] 'The Humean Theory of Motivation', *Mind* 96:36–61.
- [SOBER, ELLIOTT (2005)] 'Evolutionary Theory and the Reality of Macro Probabilities', draft manuscript (also PSA2004 Presidential Address). Currently available at: <http://philosophy.wisc.edu/sober/papers.htm>
- [STRAWSON, GALEN (1994)] *Mental Reality*, Cambridge, MA, MIT Press.
- [TOOLEY, MICHAEL (1987)] *Causation, a Realist Approach*, Oxford, Clarendon Press.
- [VAN FRAASSEN, BAS (1984)] 'Belief and the Will', *Journal of Philosophy* 81:235–256.
- [VON MISES, RICHARD (1957)] *Probability, Statistics and Truth*, revised English edition, New York, Macmillan.
- [WILLIAMSON, TIMOTHY (2000)] *Knowledge and Its Limits*, Oxford, Oxford University Press.
- [ZYNDA, LYLE (2000)] 'Representation Theorems and Realism about Degrees of Belief', *Philosophy of Science* 67:45–69.

LINA ERIKSSON AND ALAN HÁJEK  
 Research School of Social Sciences  
 Australian National University  
 Canberra, ACT 0200, Australia  
[lina.eriksson@anu.edu.au](mailto:lina.eriksson@anu.edu.au)  
[alanh@coombs.anu.edu.au](mailto:alanh@coombs.anu.edu.au)