Rationality: An Expansive Bayesian Theory

Katherine Dormandy (née Munn)
Mansfield College
DPhil in Philosophy
Trinity Term 2012

Bayesian epistemology provides a promising framework for a theory of epistemic rationality. But the way in which this framework has been built upon thus far yields an unfortunately mechanical picture of rationality, on which rational agents are mere data crunchers who receive evidential input and spit out numeric credal output. This picture is rightly criticized, most prominently by Bas van Fraassen, for being too narrow and restrictive and thus failing to account for certain features which rationality plausibly has, such as a degree of permissiveness, and for certain unconventional rational phenomena, such as conversions. Unfortunately, van Fraassen's apt criticism of mechanistic rationality overshoots its mark in seeking to topple the entire Bayesian framework. Bayesian epistemology suffers a guilt by association with the robotic picture. This dissertation aims to restore Bayesianism from the mechanistic but often implicit assumptions which corrode it, and to rebuild, from the Bayesian foundation, an alternative picture of rationality as a property of sentient agents who are capable of understanding and mentally engaging with the objects of their credences. Along the way I account for some basic Bayesian objects such as credence and evidential input. I also accord a central role to the ability of representational experiences, largely sidelined in many Bayesian discussions, to give rise to surprising evidence. On these building blocks I develop theory of rationality, Expansive Bayesianism, which evades the criticisms launched at the robotic picture and shows that Bayesianism itself is a fruitful and powerful framework for a theory of rationality.

Table of Contents

Acknowledgments

Many thanks are due to the following people: John Hawthorne and Antony Eagle for patient and knowledgeable supervision; Yoaav Isaacs, for an introduction to Bayesian reasoning and for continual encouragement and humor; Christoph Jäger, for constant friendship and philosophical mentoring; and Lizzie Fricker and Clayton Littlejohn for thoroughly and helpfully engaging with this dissertation. Thanks to Iris Geens for extreme administrative competence and concern. For friendship and encouragement many thanks go to Shirley and Mike Dobson, Silvia Jonas, Gail Leckie, Brian Leftow, Angela Matthies, Natalja Deng, Ellen and Michael McDow, Anita Renusch, and Barbara Vetter. Thanks to David Tomkins and David Harrison for prayer, encouragement, good cheer, and wisdom; Jessica Whittle for prayer, encouragement, good counsel, and lots of her wonderful self; Michael Dormandy for his abiding faithfulness and care; my parents, Irene Nordskog and Frank Munn, for so much love and almost all the kinds of support that there are; and of course God Almighty, for everything.

Introduction

This dissertation gives a theory of *epistemic rationality*. Epistemic rationality governs the way in which agents form doxastic attitudes, such as beliefs or (see below) credences. Hence it governs the way in which an agent forms and tweaks her mental representation of reality. Epistemic rationality is distinct from *pragmatic rationality*, which governs the way in which an agent decides what actions to perform to achieve her ends.[1] For reasons that will become clear later, I'll call my theory of rationality *Expansive Bayesianism*.

Rationality has etymologically to do with *roots*: a doxastic attitude is rooted when it is held fast *by* something which enables it to flourish, where a doxastic attitude, like a plant, flourishes by growing and changing. A good root allows for growth by ensuring that the plant remains connected to the source of its nourishment. What are rational doxastic attitudes rooted in? I call the root *evidence*. On the view I defend, a change of doxastic attitude is rational only if it is an appropriate response to the agent's evidence, and a doxastic response to evidence is appropriate only if the agent *conditionalizes* on her total evidence. Conditionalization is a procedure governed by a formula, Bayes's Theorem, which will be discussed below. That conditionalization is the only rational doxastic response to evidence is the definitive claim of Bayesian epistemology, or *Bayesianism*.

Bayesianism is controversial; many reject it off the cuff as restrictive, erring against epistemic growth and change by overvaluing the stability of rootedness in evidence. They accuse it of promoting a mechanistic view of rationality on which, to be rational, an agent need merely respond to numeric evidential input by producing numeric credal output which is completely pre-determined

[1] Henceforth 'rationality' shall refer to 'epistemic rationality'.

by the input value and the doxastic attitude which the agent has when she receives the evidence. Bayesian agents, the thought is, are no more than robotic data crunchers. But such accusations are inapt. Bayesianism comes in many forms, and Expansive Bayesianism, the view developed here, has wide scope even for rational doxastic changes which are agent-directed as well as perhaps surprisingly liberal. I postulate that many non-Bayesians who see how powerful a theory Bayesianism can be when harnessed with the epistemological postures I take will thereby have good reason to become Bayesians themselves.

A prominent non-Bayesian is Bas van Fraassen. He endorses conditionalization but denies that it is a *necessary* component of rational response to evidence, and hence rejects Bayesianism as I (and he) characterize it. Van Fraassen argues that Bayesianism has no room for forms of doxastic development which involve liberal and far-reaching changes of view, such as scientific or religious conversions, and that it straitjackets epistemic agents from being rationally able to exercise the full extent of their epistemic agency. These criticisms are apt and valuable. But they do not, I argue, apply to Bayesianism itself. They apply to a particularly harsh form of it, which I'll call *Restrictive Bayesianism*. In pointing out these problems van Fraassen has performed a valuable service to epistemology, but in thinking that they touch Bayesianism itself, he is mistaken. Much of the burden of my dissertation is to distinguish Restrictive Bayesianism from my view, Expansive Bayesianism, and to show that the latter does not face the problems which van Fraassen isolates. Doing this will require knocking down some cherished but often implicit shibboleths associated with Bayesian epistemology.

Ideal Rationality

Expansive Bayesianism is a theory of how an ideally rational agent integrates new evidence into her doxastic system; in other words, it is a theory of ideally rational change of doxastic attitudes over time. Note the modifier 'ideally'. There are two views of rationality, which might at first appear to conflict. Both views take rationality to be a source of norms for the formation of doxastic attitudes, but on one view, which we'll call the *ideal* view, rationality designates a form of doxastic perfection which human agents can rarely if ever meet. According to what we'll call the *regulative* view, by contrast, rationality designates the way in which imperfect, everyday agents ought to form their doxastic attitudes. An ideal agent always starts off probabilistically coherent (on which see Chapter 2) and thus is always in a position to conditionalize on receipt of evidence. One way of being non-ideal, by contrast, is to start off incoherent – since conditionalization is therefore not an option for her, ideal rationality offers her no advice, and regulative rationality must step in to tell her how to

respond to her imperfect situation. Ideal rationality governs agents who never make mistakes; regulative rationality governs agents who already have.

Ideal and regulative rationality are compatible, however, and I think that a complete theory of rationality will treat both: rationality does set standards the perfect achievement of which is impossible for human beings, but there are also more and less rational ways for agents who do not meet these standards to form and manage their doxastic attitudes. I think that rationality is an ideal which admits of degree, and that a complete theory of rationality will characterize both the perfect as well as the imperfect instances, and account for whatever closeness relations there are between them.

But Expansive Bayesianism (indeed, as we'll see below, Bayesianism full stop) pertains to ideal rationality only. I offer no guidelines for imperfect agents. A theory of rationality has to start somewhere, and I hope that my insights into ideal rationality can provide a springboard for a theory of regulative rationality – at least assuming, as I do (but recognize that many do not), that regulative rationality somehow takes its cue from the ideal.[2]

Why is a Theory of Rationality Important?

I am interested in the rational doxastic response to evidence because it is the locus of several important controversies in epistemology. One pertains to scientific theory change: what if anything makes a scientist (or any agent, for that matter) rational in changing from taking one theory to be true or probable to endorsing another instead? This question is particularly interesting when the theory in question is large-scale and foundational, such that rejecting it would require much epistemic soul-searching. Can evidence ever make such extreme forms of theory change rational? And if not, were the historically significant changes of this sort rational? An advantage of Expansive Bayesianism over Restrictive Bayesianism is that it can show how even radical scientific (and for that matter, non-scientific) conversions can be rational.[3]

Another area impacted by a theory of the rational response to evidence is the newly burgeoning debate about the epistemology of disagreement. This debate focuses on the question: to what extent if any does the disagreement of someone whose opinion an agent respects provide evidence, for that agent, against her own view? This question has been extensively dealt with by appeal to

[2] See Foley (1993) for the view that the whole project of constructing unattainable rational ideals is misguided, and Christensen (2004), Chapter 6, for a sensible take on the relationship between regulative and ideal rationality.

[3] Chapter 6 discusses this issue.

thought experiments on which readers' intuitions must rule in one way or another.[4] But without a framework for understanding what evidence is and what makes for a rational response to it, the theoretical power of such approaches is bound to be limited. Insight into the epistemology of disagreement must stand on the shoulders of a theory of the rational response to evidence. Though I do not discuss disagreement in particular, I take Expansive Bayesianism to provide a crucial theoretical backdrop for pursuing some of its fundamental issues.

A much more venerable controversy than disagreement – or perhaps the classic example of it – surrounds how and whether religious faith can be rational. Central to this question is the important debate which began in antiquity yet rages on today, over what role if any experience, understood in opposition to argumentation, plays in making an agent rational in adopting (or maintaining) the doxastic attitudes which accompany her faith. Views range from claiming that experience has little if any relevance to claiming that it is the *sine qua non* of reason for faith, with various nuanced views in between.[5] By and large the participants in this debate operate on the presupposition that there is experience on the one hand, and something which they call 'evidence' on the other hand. Expansive Bayesianism exposes this assumption for the falsehood it is, presenting a simple characterization of evidence according to which any experience which provides epistemic reason thereby provides evidence. And by the same mechanism which Expansive Bayesianism shows that scientific conversions can be rational, it shows how religious conversions can be rational too.

Credences

Here is the place to jettison the cumbersome terminology of 'doxastic attitudes' and replace it with something sharper. 'Belief' is an option, but not the focus of this study; the reason will shortly become clear. I am interested instead in credences. 'Credence' is the name for the doxastic attitude featuring in *fine-grained epistemology*, that is, epistemology which takes there to be doxastic attitudes which admit of degree.

A *credence* in a proposition P, as I understand it here,[6] is a doxastic attitude which expresses *how probable* an agent takes P to be, by assigning a precise point-valued number to P.[7] This is different from a belief, which merely expresses a more general and unquantified endorsement of P. It is also different from the belief that P is probable to such-and-such a degree. And it is different from a

[4] For a variety of takes on disagreement see Feldman and Warfield (2010), Elga (2007), Christensen (2007), Kelly (2005), and Lackey (2008).
[5] For some canonical contributions to a vast literature see Swinburne (2004), Plantinga (2000), and Mackie (1982).
[6] I follow Eriksson and Hájek (2007) in taking credences to be unanalyzable; more in Chapters 1 and 5.
[7] I assume that propositions are the objects of doxastic attitudes; we may think of them generally as the complementizer phrases that appear at the end of 'that'-clauses.

credal attitude which expresses how probable an agent takes P to be yet which does not assign a number to P.[8] Instead, a credence is the expression of what we may think of as a 'degree of confidence': the attitude *itself* admits of degree and hence is quantifiable; an agent has a credence of some number *x* in a proposition P.[9]

The notion of 'degree' already commits us to particular assumptions about the way in which doxastic attitudes are quantifiable,[10] on which more shortly. But the less precise notion of 'level of confidence' already plays an important but undervalued role in epistemology. An important question has traditionally been whether an agent is rational in adopting a belief that P. But recognizing that there are also doxastic attitudes which admit of degree opens up an entirely new dimension of epistemic evaluation: namely, whether an agent is rational in having some level of confidence or other that P. A theory of evidence which accounts for rational *belief* only accounts for part of the picture. For surely, if evidence can justify thinking that a proposition is true, it can also justify levels of confidence.

Indeed, much epistemology (including the three areas I mentioned above, scientific reasoning, disagreement, and the faith/reason controversy) already incorporates notions of 'levels of confidence'[11]. Moreover, degrees of confidence can cast much light on the overlap between epistemic and pragmatic rationality. An agent may be somewhat confident that it will rain yet not bother to take his umbrella because lugging it around is a major hassle, and he may anyway think that, in spite of his confidence, there is still a decent chance that it won't rain. But an agent who is *highly confident* that it will rain (and who wants to stay dry) will take it, in spite of the hassle of doing so. Although both agents may count as believing that it will rain, distinguishing between how strongly they believe this – what level of confidence they have – helps explain their differing behavior. Finally, much epistemology, particularly in confirmation theory and the philosophy of science more broadly, already operates with degrees of confidence. I work from the assumption that the notion of credence can stand in as a handy and illuminating precisification of talk of levels or degrees of confidence.

[8] See e.g. Sturgeon (2011) and Hawthorne (2009).

[9] There are important objections to the project of quantifying doxastic attitudes, especially to quantifying them with point-valued measurements. One is that agents have no introspective access to such precise values, and another is that doxastic attitudes themselves do not admit of such precisification. For discussion of these and a few other objections, see Christensen (2004).

[10] Some take credences to be 'mushy', in the sense of spanning multiple points over the unit interval; for such a view see van Fraassen (1995), and for arguments against it see White (2009) and Elga (2010). For simplicity I will ignore such views and adopt the idealizing assumption that credences always have precise values. What I say in these pages might be extended to accommodate 'mushy' credence, but I won't try here.

[11] See e.g. the above references to the disagreement literature.

Credences make such talk precise because they can be mapped one-to-one onto the unit interval, with credence 1 as certainty and credence zero as the opposite.[12] This enables credences to be governed by the probability calculus, which gives axioms ranging over point-valued units between zero and one. Credences are expressions of *subjective probability values*, which is to say, probability values endorsed by an epistemic subject.[13] The relationship between probability and doxastic attitudes opens up tantalizing avenues for approaches to epistemic rationality. With good reason, the most developed approach is *Bayesian*, by which I mean that it holds that conditionalization, a rule furnished by Bayes's Theorem of the probability calculus, is the only (ideally) rational way to change one's credences in response to evidence.

Bayesianism

Bayesianism is not theory of rationality. It is a set of axioms and theorems about probability functions, where these functions are understood to model rational credence. Bayesianism thus provides a *framework* for a theory of rationality. By this I mean that the probability axioms place constraints on the sorts of relationship that may rationally obtain among credences, both at a time and from one time to another, but that, on a vast array of fundamental epistemological and metaphysical questions, Bayesianism is silent. It says nothing about the most basic properties of credences, about what evidence is, what rationality is (beyond adherence to the Bayesian axioms), what individuates epistemic agents, what is needed to be rational from one time to another (beyond conditionalization on one's evidence), what the object of credence is, to list just a few.

Aside from the minimalism both of Bayesianism's constraints on rationality and of its commitments about the items in the epistemological domain, there is another fashion in which Bayesianism is minimalist. To appreciate this second way of being minimalist, recall my focus on *ideal* as opposed to regulative rationality, where ideal rationality postulates an epistemic perfection which human beings will never meet, and regulative rationality gives norms for everyday human agents. If we think of a theory of (ideal) rationality as a house that is furnished and ready to live in, we may think of Bayesianism as the cement foundation plus the wooden skeleton, but without walls, plumbing, flooring, and so forth. Bayesianism does not give a complete rational ideal, but it provides the framework for one.

Nor, importantly, is Bayesianism regulative: through and through it is a framework for a rational ideal. Hence it does not rule on what standards of rationality if any govern agents who fall short (but

[12] Thus one assumption I am making, which may be contested, is that credence 1 is the same as certainty.
[13] Sometimes in the following I use 'credence' loosely as the probability value itself, rather than the attitude which expresses this value.

perhaps not hopelessly short) of ideal rationality. It does not tell us how to take a dilapidated house and re-model it into a Bayesian one. If agents have failed to change their credences by conditionalizing, they are out of the Bayesian jurisdiction; if their credences do not conform to the axioms of probability, Bayesianism tells them to resolve this situation immediately but it does not tell them in which of infinitely many ways to do so. This is a deep and important limitation, for rationality is normative over imperfect agents if it is normative over any agents, yet it is precisely the imperfect agents who fall out of the Bayesian purview. This limitation must eventually be addressed if Bayesianism is to be shown to offer a framework for a complete theory of rationality, including both ideal and regulative rationality. But my focus here is on making it a viable ideal, not extending it to govern non-ideal agents.

Bayesian Minimalism

There are thus three ways in which Bayesianism is minimalist: (1) its constraints offer the framework for a theory of ideal rationality, without itself providing such an ideal; (2) it is committed to few assumptions about the nature of things falling under epistemological categories; (3) it does not touch on regulative rationality at all.

Let us leave aside the third way of being minimalist and focus on the first two. On the one hand, it is precisely Bayesianism's minimalism in these two ways which makes it so potentially fruitful. Bayesianism can lend itself to highly productive epistemology by providing some beacons for orientation but otherwise leaving the surrounding terrain to be developed in any number of ways. On the other hand, this minimalism also poses a danger for Bayesian approaches: It makes them vulnerable to being overrun by assumptions which quietly and systematically infiltrate the theoretical gaps which Bayesianism leaves open. Too often these assumptions are not discussed explicitly and risk being mistaken for Bayesianism itself, just as vines, when thick enough, risk being mistaken for the house or the tree which they strangle. An example of an assumption which I discuss at some length (Chapters 5, 6, and 7) pertains to the nature of conditionalization itself, and turns out to have far reaching implications for an epistemology built around it, since it restricts the way in which a response to evidence can count as a conditionalizing response to begin with – and recall that conditionalization is the only rational form which such a response can take.

Sometimes these assumptions show off Bayesianism's best form, but most often – especially when they grow unsystematically as 'default' assumptions which serve the purposes of those using Bayesianism in one specialized context or another – they limit the potential of a theory of rationality which is built on it to be the best such theory it can be. A theory of rationality should say of all ideally

rational credence changes that they are rational, and of all irrational ones that they are irrational. Van Fraassen criticizes Bayesianism for failing to do this. He says that there are some ideally rational updates which Bayesianism counts as irrational. My contention is that it is not *Bayesianism* which he fingers but rather what I call 'Restrictive Bayesianism', which is a Bayesian theory of rationality which builds in false assumptions.

My goal in this dissertation is thus to untangle and uproot some of these implicit default assumptions which have grown haphazardly around the Bayesian shell. In particular I will isolate a few which will handicap Bayesian attempts to model rational conversions or to allow for epistemic agency. And in areas where no particular assumptions have taken hold but diverse ones vie for dominance, I argue for one or another. In short, this dissertation is an attempt to perform some much needed gardening work around the Bayesian foundation, tearing up non-Bayesian assumptions and replacing them with properly cultivated claims, as well as shearing the parasites which threaten to overrun any of the claims which have attached themselves to the edifice yet deserve to remain, and to present both a true and insightful Bayesian theory of rationality.

My theory, Expansive Bayesianism, does two things. First, it fills in some areas which Bayesianism leaves open, such as those pertaining to the nature of evidence, credence, and so forth. Second, it makes explicit and then rejects some parasitic default assumptions which are typically, but implicitly, associated with Bayesianism. In company with the right theoretical additions, and shorn of the wrong ones, Bayesianism is a powerful and flexible framework for a theory of rationality, and Expansive Bayesianism intends to show it off at its finest.

Chapter by Chapter Summary

Chapters 1-3 lay out some important initial groundwork which will prepare the way for the core of Expansive Bayesianism to be presented and applied in **Chapters 4-7**. The rest of this introduction gives a chapter-by-chapter narrative.

The first task, performed in **Chapter 1**, is to map out the domain of fine-grained epistemology (the epistemology of credences). In particular, I distinguish rationality from a few other salient sources of epistemic goodness. This needs to be done, first, so that rationality can be cordoned off for examination within its proper context, and so that it is not mistaken for any of the other sources of epistemic goodness. Second, since systematic fine-grained epistemology is in its infancy, so are discussions of the sources of epistemic value within it. It is particularly important to obtain a handle on fine-grained epistemic categories, because many of them differ significantly from coarse-grained categories, just as credence differs from belief. An example is that beliefs can have the epistemically

good-making property of *truth* predicated of them, but credences cannot, since they admit of degree whereas truth does not. Yet there is an analogue of truth which does confer a fine-grained form of goodness: proximity to certainty in truths. This, plus rationality and propositional grasping, are the three salient sources of fine-grained epistemic goodness which Chapter 1 isolates. Subsequent chapters narrow in on rationality.

But propositional grasping will also play an important role, as we'll see, though it is not typically acknowledged – either in fine- or coarse-grained epistemology – as counting among the sources of epistemic goodness. In developing my theory of rationality, I hope to correct this distortion. Chapter 1 sets the stage by arguing that propositional grasping is an epistemic good which is logically independent of rationality: an ideally rational agent may fail to grasp some propositions, and an agent who grasps all propositions may fail to be rational.

Chapter 2 characterizes Bayesianism as providing set of necessary conditions on rationality: hence as the framework for a theory. One condition constrains rationality at a single point in time, and another constrains rational change of credence from one time to another. The condition on synchronic rationality is

> An agent is rational at a time only if his credences are a probabilistically coherent function. (*Bayesian Statics*),

and the condition on diachronic rational credence is

> An agent is rational from one time to another only if, when she updates on evidence, she does so by conditionalizing. (*Bayesian Updating*)

I distance myself from those Bayesians who add the claim that these two conditions are also sufficient for rationality at a time and over time: one of the first of the non-Bayesian weeds which I uproot.

Chapter 3, which has three parts, begins building on the Bayesian foundation by accounting for some of the key elements in fine-grained epistemology, credence and evidential input.

Part I of Chapter 3 highlights another assumption with which Bayesianism is sometimes but needlessly associated and which I reject, namely:

> Rational agents have credences defined over all propositions. (*Complete*)

Complete is often assumed because the probability axioms themselves are often given with scope over all propositions. But since having a credence in a proposition entails that the agent grasp that

proposition (so Chapter 3 argues), Complete entails that grasping all propositions is necessary for rationality. But propositional grasping, Chapter 1 will have argued, is an ideal epistemic property which is distinct from rationality, and logically unrelated to it (except insofar as a single credence requires its object to be grasped).

I thus deny Complete, arguing that the probability axioms do not essentially quantify over all propositions. I recommend applying the probability axioms in epistemology in a manner that restricts them to the set of propositions over which an agent's credences happen to be defined. This move reflects the fact that even ideally rational agents can fail to grasp some propositions and thus fail to have credences in them. And this, in turn, allows me to account, in later chapters, for the way in which agents can undergo rational credence change which they could never have expected: by coming to grasp a new proposition and forming completely new credences conditional on it. We'll return to this in due course.

Part II of Chapter 3 continues to set the stage for Expansive Bayesianism by giving an account of some of the most basic elements in the Bayesian domain, including credence and evidence. These claims work toward a picture of rationality as a property of sentient agents capable of entertaining and engaging with the objects of their credence, as opposed to a mechanical picture of rational agents as machines programmed to crunch numeric data.

I establish some Expansive Bayesian claims about evidence and input, calling much needed attention to the logically first step in updating, in which an agent assigns a credence to her evidence (as opposed to the second, which is the conditionalizing step). Those Bayesians who do explicitly address the input step do so by placing rational constraints on it at exactly the wrong places, I'll argue. Typically there is a constraint on the type of representational experience on which the input step can be based; I argue that there is no such constraint. Typically, also, Bayesians hold either that there is no constraint on the particular proposition which an evidential experience makes salient as evidence, or that this proposition may only pertain to the way things seem to an agent (as opposed to the way that they are); and they hold that there is no constraint whatever on the numeric value which the agent assigns to her evidence. I argue that that all of these Bayesian constraints are flawed, and that there are constraints on both of these things, though not the constraints typically endorsed by Bayesians.

Finally, Part III of Chapter 3 draws out some of the large-scale implications of this account of credence and input. It argues that the coming to grasp of a new proposition is a potential source of evidence which can, depending on the input credence which the agent assigns to it, radically alter an agent's representation of reality, and that coming to grasp new propositions is a much more regular occurrence than one might think. These insights pave the way for the following chapters to begin

making my case that Bayesianism is a flexible and powerful theory of updating when divested of certain burdensome assumptions with which it is often, unnecessarily, associated.

On account of some of these assumptions, some argue that Bayesianism is an overly restrictive theory which cannot accommodate some important rational phenomena. A prominent critic along these lines is Bas van Fraassen. **Chapter 4** presents the theory which van Fraassen advances in lieu of Bayesianism, which he calls 'voluntarism'. Voluntarism, argues van Fraassen, gets some things right which, he claims (wrongly, as it turns out), Bayesianism cannot. Chapter 4 contents itself with laying out voluntarism and discussing some problems internal to it. The upshot is that these problems – including, prominently, van Fraassen's endorsement of a strong variety of doxastic voluntarism which I argue is false – speak strongly against voluntarism, regardless of whether van Fraassen's critiques of Bayesianism are on the money.

Fortunately, however, they are not. **Chapters 5, 6, and 7** reveal that van Fraassen's critiques are levelled against a form of Bayesianism which is overly restrictive and hence severely hobbles the powerful and simple claim that conditionalization is a necessary component of any rational response to evidence. These chapters present and apply the bulk of the Expansive Bayesian account of rational credence change, showing that Expansive Bayesianism evades some prominent critiques launched against an unduly strict form of Bayesianism which has until now tended to be mistaken for Bayesianism itself.

On the shoulders of the claims advanced in Chapters 1 through 3, **Chapter 5** introduces the way in which Expansive Bayesianism models the rational doxastic response to evidence. It does so obliquely, by calling attention to an assumption with which Bayesianism is traditionally associated, and arguing that Bayesianism is not committed to it at all. The assumption is:

> Once an agent's credence in a proposition *p* has reached 1 (and hence her credence in not-*p* has reached 0), any change of credence in *p* (or in not-*p*) is irrational. (*Unrevisable*)

An argument which might be offered for Unrevisable depends on the premise that

> No conditionalizing update can move a credence down from 1 or up from 0. (*Stuck*)

But Stuck is false. Conditionalizing updates, I go on to show, can result in a credence of zero moving above zero, and a credence of 1 moving lower than 1. It becomes clear that they can do this when we notice that conditionalization proceeds in two steps: the *input step*, in which the agent assigns a credence to her evidence proposition, and the *adjustment step*, in which she sets her posterior

credences equal to her prior credences conditional on her evidence (or performs the analogous manoeuvre for uncertain evidence).

There is a habit of mind among Bayesian epistemologists which, I suggest, makes it easy to slip into some of the particularly noxious assumptions which, when conjoined with Bayesianism, siphon much of its power and flexibility. The habit of mind is the tendency to give undue focus to the second step of conditionalization, the adjustment step (where Bayes's Theorem is applied), at the expense of the first, the input step. It is easy to imagine how this distortion would arise, given the excitement over the way in which the adjustment step proceeds neatly by formula once input values are specified, whereas the input step requires values which the agent apparently plucks subjectively from her experience.

In ignoring the input step, I argue, Bayesian epistemologists have tended to propagate a false assumption which entails Stuck, about the nature of conditionalization. The assumption in question makes trouble for Bayesian attempts to respond to two criticisms launched by van Fraassen, which Chapters 6 and 7 address. The assumption is this:

> A *conditionalizing update* is one in which the agent sets her new credences equal to the value of the credences, conditional on her new evidence, which she has before she receives evidence. (*Pre-Input*)

Pre-Input limits the credence which the agent may rationally assign to her evidence in the input step. It is because of what Pre-Input assumes about how conditionalization works, not because of anything inherent to conditionalization itself, that Stuck and hence Unrevisable are also thought to be true. When we reject Pre-Input, which nothing about the probability axioms or Bayes's Theorem forbids us from doing, we see that conditionalizing updates can bring an agent's credence up from zero or down from 1, and thus that Unrevisable is false. Already we are starting to see how Bayesianism provides a flexible and powerful framework for liberal credence change which is nonetheless rooted in evidence.

Chapter 6 addresses one of van Fraassen's reasons for rejecting Bayesianism, which is that the latter is allegedly unable to model conversions, such as scientific or religious ones. Bayesianism, says van Fraassen, entails that such changes of credence are irrational – which he patently denies. I agree with van Fraassen that conversions can be rational. He has done a great service to epistemology in highlighting their importance and in arguing that no theory of rationality is correct which cannot at least provide an adequate model of them. But van Fraassen is mistaken in thinking that Bayesianism cannot provide such a model. Expansive Bayesianism can model them easily. Moreover, unlike van Fraassen's account of rational conversions, which locates their rationality in their stemming from the

agent's *will*, Expansive Bayesianism gives a much more plausible account of them, as coming about because agents receive evidence which they could not have anticipated. Conversions, far from being powered by the agent's decision to adopt this credence or that, are just rational responses to evidence. Thus Expansive Bayesianism not only provides an adequate account of conversions; it sheds much needed insight on their nature.

Chapter 7 discusses a view, which I call *Permissive Updating*, which says that rationality sometimes permits the assignment of more than one posterior credence to a proposition on receipt of a given piece of evidence. Expansive Bayesianism, I argue, provides a counterexample to the claim that Bayesianism is incompatible with Permissive Updating. Permissive Updating has a prima facie appeal for several reasons which I argue are illusory, and hence is a bad claim on which to stake the fate of Bayesianism. But anyway, Expansive Bayesianism is perfectly compatible with Permissive Updating. In particular, Expansive Bayesianism shows there to be scenarios in which rationality may well permit the assignment of more than one posterior credence.

Conclusion

Expansive Bayesianism is a theory of rationality built on the Bayesian foundation. It seeks to preserve and adorn the minimalist beauty of this foundation by adding features which show it off at its best and by removing those which have grown around it haphazardly and threaten its structural integrity or at least disable it from supporting more development. It pioneers a map of the fine-grained epistemological territory by postulating the epistemic goods of rationality, certainty in truths, and propositional grasping. It endorses the definitive Bayesian claim that conditionalization is necessary for any rational response to evidence, thus rooting rationality firmly in evidence. But it construes evidence as broadly as possible, to incorporate anything, including intellectual seemings, which count as epistemic reasons. It further argues that new evidence can, often surprisingly, be had by receiving as evidence propositions which, before receiving them, were doxastically impossible, or by coming to grasp new propositions, which, I argue, happens more often than many think. It then shows that Bayesianism, when coupled with these claims, is compatible with many perhaps surprisingly liberal changes of credence, *pace* van Fraassen: it is compatible with epistemic impossibilities becoming possible, epistemic certainties becoming uncertain, and with the claim that there is sometimes more than one rational credence to adopt on the basis of any given piece of evidence. In short, anyone who rejects Bayesianism because he thinks that it is excessively strict is probably thinking not of Bayesianism but of some non-Bayesian assumptions often coupled with

Bayesianism. Cutting these assumptions away makes room for Expansive Bayesianism, which maintains epistemic leniency while allowing for a healthy evidentialism to grow in their place.

This is not to deny that many issues remain which I do not address. My focus is on using Bayesianism to model rational credence formation generally, not to deploy it (for example) in solving particular confirmation-theoretic problems. Hence I leave relatively untouched questions about how Bayesianism should recommend updating in this or that precise situation (such as those raised by the problem of old evidence, Sleeping Beauty cases, and so forth). I am aware that some of these types of situation are challenging enough to vitiate the whole Bayesian project, or at least to show its application to be significantly limited. But I hope that Expansive Bayesianism will provide a promising framework within which some or many of these issues can be addressed. The proof will be in the pudding, though, and the pudding is still a work in progress. In short, my aim is to make explicit issues which should but have not yet received the philosophical attention which they ought if Bayesianism is to grow to maturity and then prove its mettle as a framework for a complete theory of rationality.

1. Epistemic Goods and Terms of Discussion

Introduction

Before beginning it is important to clarify some of the terms under discussion and give an overview of the conceptual territory in which I am operating. This chapter does these things. It begins by characterizing the ways I am using the terms of interest, including *epistemic agent*, *credence*, *object of credence, proposition*, and so forth. *Rationality* of course will be characterized by this dissertation as a whole. But rationality is only one of several sources of epistemic goodness. This chapter will elucidate two other important sources: being certain in true propositions, and grasping propositions. Acquaintance with these epistemic goods helps us to see what rationality is not, and provides perspective on what it is. The bricks laid in this chapter will put us in a position to start, in the following chapters, the construction of a theory of rationality itself.

Terms of Discussion

I'll begin by characterizing, as much as is needed for my purposes, some of the key notions with which this dissertation will deal.

Epistemic Agent

An *epistemic agent* is the subject who has doxastic attitudes (such as credences or beliefs). The prototypical epistemic agent is the human being, but an agent is any being whose cognitive apparatus is sophisticated enough to form a representation of reality (for discussion of this notion see below), such as angels or God. I am assuming that having such a cognitive apparatus includes being sentient and having the ability to entertain the objects of one's doxastic attitudes. In this way my view differs from a typical Bayesian view according to which epistemic agents may as well be numeric data crunchers who have no awareness of what they are doing. On my view, all persons are epistemic agents (though perhaps non-persons, such as some animals, can be too). Moreover, when a person is an epistemic agent, then the person and that agent are identical.[14] Thus being an epistemic agent is not like being a legal agent, since some persons, such as children, are not legal agents. It is more like being the agent of an action: even if not legally responsible for their actions, children perform them. Similarly, children form credences, and hence are epistemic agents.

An *ideal epistemic agent* is one who perfectly possesses some form of epistemic good. There are three salient such goods which we'll discuss below, of which rationality is one. Recall from the introduction that this thesis gives a theory of ideal, as opposed to regulative, rationality; it thus characterizes a form of epistemic perfection which human agents rarely if ever meet. Such a theory is intrinsically interesting but also, hopefully, of some use to those who aspire to set normative standards for imperfect agents. Even if we cannot be perfect like (say) God is perfect, we can look to his characteristics as a model to imitate in one way or another. My approach is thus to be distinguished from those which suppose that ideals are of little relevance to non-ideal agents.[15] Ideals, I think, can be ideals even for agents who can seldom if ever embody them. But my goal here is neither to argue for this claim nor to show how the rational ideal I present can be normative over non-ideal agents. I only characterize that ideal itself.

Later on this chapter characterizes three forms of epistemically ideal agent (one for each form of epistemic good), which are important to distinguish so as to set in relief one of these goods, rationality.

[14] What about beings whose cognitive faculties are somehow abnormally fragmented, like people who suffer from severe memory loss or multiple personality disorder? If there is an extent to which a person can be fragmented, then so, to that same extent, can an epistemic agent be fragmented. But since I am dealing here with ideal epistemic agents, the question of fragmentation will not arise.

[15] For such an approach concerning rationally ideal agents, see e.g. Foley (1993).

Credences

The epistemic agents under consideration here have credences. We saw already that a credence is a kind of doxastic attitude, where a *doxastic attitude* is a mental state which is intentionally directed 'toward' some piece of representational content or which is 'about' that content[16] (I am calling such content 'propositions', on which more shortly). A credence is a doxastic attitude which expresses an agent's subjective probability value for a proposition.

The Introduction noted that I take credences to be unanalyzable. Thus they are not defined in terms of ways in which agents bet (*pace* de Finnetti, 1990) or of dispositions to bet in particular circumstances (*pace* Jeffrey, 1994).[17] Nor, in keeping with my methodological assumption that fine-grained categories are *sui generis*, are credences defined in terms of beliefs: a credence of *x* in *p* is not the belief that *p* is probable to degree *x*. Among other problems with this suggestion, it makes it a requirement of having a credence that the agent be competent to deploy the concept of credence; but surely an agent can use her credences to bet (which is a sign that she has them) without possessing this concept.[18]

How are credences individuated? From the use of 'credence' in the literature, one might think that there are two competing suggestions. Sometimes credences are spoken of as if they are a doxastic attitude which endures through changes of numeric value, as in: 'My credence at *t1* was ½ but at *t2* it was ¾'. This usage would suggest that the numeric value is distinct from the attitude itself, which is individuated by the proposition in which it is a credence. But sometimes credences are spoken of as if they *are* the numeric value, as in: 'on receipt of evidence she acquired a new credence in P'. The latter usage designates a probability function, where such functions of course do change whenever their value changes.[19] How are credences to be individuated: by the propositions which are their objects or by the numeric values assigned to those propositions? I think that credences are individuated by their numeric values. The uses of 'credence' which seem to refer to a doxastic attitude independent of the number can be explained by noting that 'my credence' is a description (short for 'the credence of KM'). Like any description, it can change its denotation from one context to another; hence 'my credence in *p* at t' need not refer to the same credence as 'my credence in *p* at t1'.

[16] This usage is in keeping with the tradition of Brentano (1874) and Husserl (1900).
[17] For a good criticism of these views and good discussion and defense of the claim that credences are unanalyzable, see Eriksson and Hájek (2007).
[18] Can beliefs be reductively defined in terms of credences? I don't think so, but won't discuss the issue here. See Sturgeon (2008 and 2011), Foley (2009), and Frankish (2009).
[19] The grammar of 'credence' is also in its infancy: sometimes it is used as a count noun, as in 'she had a credence in each of ten propositions', and sometimes as a mass noun, as in 'she has firm credence'. I employ both.

That credences are both doxastic attitudes and probability functions entails that agents do not have credences unless their fine-grained doxastic attitudes are subject to measurement by the probability calculus, which entails in turn that they obey the three probability axioms (see Chapter 2). Since obeying these axioms is a component of rational perfection, no agents can have credences unless they are rationally ideal. Already this restriction seems to remove Bayesian epistemology one step away from application to non-ideal agents. But this conclusion is too quick. For real agents have fine-grained doxastic attitudes even if these attitudes, strictly speaking, are not credences. Either 'credence' can simply acquire a broader meaning, to refer to fine-grained doxastic attitudes that are not probability functions, or another word can be coined to refer to these non-ideal attitudes. Since I am focusing on ideal agents, 'credence' suits my purposes.

Credences come in two varieties: unconditional, and conditional. An unconditional credence of *x* in *p* is an expression of the agent's degree of confidence, namely *x*, in the proposition which is the object of her credence, namely *p*. A conditional credence, likewise, is an expression, but it is an expression of the degree of confidence that the agent has *on the supposition* that he becomes certain of some other proposition. Thus a credence of *y* in *p* conditional on *q* is the degree of confidence, *y*, which the agent has on the supposition that he becomes certain that *q*. A conditional credence is thus an expression of the agent's assignment of subjective probability to one proposition relative to another. The notation for a credence in *p* conditional on *q* with a value of *y* is 'Cr ($p \mid q$) = *y*'. Chapter 2 says more about how unconditional and conditional credences relate.

The Objects of Credence

'Proposition' is the term I use to denote the objects of credence, and of doxastic attitudes more generally. I do not claim to give a full account of propositions, though I will make some initial remarks about the kind of semantic object which is suitable for performing the theoretical function of the object of credence. First, propositions can be true or false. Second, they can be expressed, relative to contexts, by declarative sentences which are uttered assertorically or which appear as the complementizer phrase that comes at the end of a 'that' clause, as in 'She has a high credence that *the Yankees will win the World Series*'. Propositions thus function as the semantic values of declarative sentences. Third, a salient and fairly uncontroversial feature of propositions is that they carry representational content; that is, they represent the world as being a certain way (more on this shortly).

Fourth, I construe propositions as having logical form: it makes sense to talk about some propositions as being tautologies and others as being contradictions. Although this is a common

assumption, it is not completely uncontroversial. For example, some think of propositions as sets of possible worlds. On this picture, propositions do not have logical form, and any logical properties are ascribed to sentences rather than to propositions. By contrast, the picture I am working with – which is fairly standard – says that logical properties can correctly be ascribed to them. Note, importantly, that on this view the notion of a logical contradiction is different from that of a metaphysical impossibility, and the notion of a tautology is different from that of a metaphysical necessity. It may be impossible for an avocado to compose a Viennese waltz, but there is no logical contradiction in the claim that it can – that is, the claim cannot be reckoned false simply on account of its logical form. Similarly, it may be necessary for an avocado to fail to compose a waltz, but there is no tautology here: the claim that an avocado cannot compose a waltz cannot be identified as true merely by examining its logical form.[20] Thus, to take a well known example, when 'Hesperus' denotes the same item as 'Phosphorus', my view makes the fairly standard assumption that, although both 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' are metaphysical necessities, only the former is a tautology. If talk of tautologies is going to apply to propositions, the proposition expressed by 'Hesperus is Hesperus' has to have more structure than simply the relevant object and the identity relation – 'Hesperus' must contribute something like a sense as well as a referent.

Fifth, what makes one proposition different from another? What I have said already imposes certain constraints. That proposition A and B are true in the same metaphysically possible worlds does not suffice for their identity, since one may be a tautology while the other is merely metaphysically necessary. Thus my picture is that propositions are more finely individuated than sets of possible worlds. Of course, this leaves many issues of grain and individuation unresolved, but this work cannot undertake a complete exploration of the metaphysics of propositions. In general, the more coarsely propositions are individuated, the fewer sorts of content they can represent. Since doxastic attitudes can be had toward very finely individuated pieces of content, I take it that the more finely propositions are individuated, the better they can serve in a model of doxastic attitudes. Hence propositions, as I am thinking of them, have enough structure to represent very fine gradations of content.

Note that in a setting of fine-grained epistemology it is easy to define up a notion of which sets of propositions are doxastically possible: they will be those sets to the conjunction of which an agent assigns non-zero credence. A proposition is *doxastically possible* for an agent just in case it is not excluded by any proposition of which she is certain.[21, 22]

---

[20] .

[21] 'Epistemic possibility' is sometimes used casually in this sense. But I am using the more accurate term 'doxastic possibility' in order to distance myself from a more robust and technical notion of an epistemically possible proposition as one which is not excluded by any proposition which the agent *knows*. Since knowledge

A corollary is that a set of propositions is only doxastically possible if the agent can grasp them, on the assumption (which Chapter 3 defends) that one can only assign a credence to a proposition which one grasps, or "can think" (Williamson, 2000, p. 194). Grasping a proposition *p* does not require being competent in any language; I take it that non-linguistic animals can grasp propositions, such as *There is food before me*.[23] Having an experience as of *p* entails grasping *p*. Grasping as I refer to it here is a coarse-grained propositional attitude: either an agent grasps a proposition or she does not – I do not work with degrees or levels of grasping. But my account could easily make room for a notion of finer-grained propositional grasping.[24] It is for simplicity that I deal only with the coarse-grained variety.

My account of the objects of credence might be thought to raise trouble for a Bayesian framework for rational credence, since Bayesianism requires that agents have credence 1 in tautological propositions. I will address this worry in Chapter 2, after introducing and discussing the Bayesian requirements for rational credence.

This completes my exposition of some of the main terms of discussion. Others will be characterized when they are introduced. I'll now say a few words about the notion of epistemic goodness, and some particular epistemic goods we will encounter in these pages.

Three Epistemic Goods

An *epistemic good* is a property the possession of which confers goodness of a uniquely epistemic kind. I will not attempt to define the notion *epistemic*. But we may think of it most generally as pertaining to the way in which the agent mentally represents reality.[25]

Mainstream epistemology takes there to be two mutually exclusive and exhaustive kinds of epistemic value, internalist and externalist. A form of epistemic value is *internalist* just in case it is

---

that *p* entails that *p* is true, then, on this way of thinking, a proposition is epistemically possible only if consistent with truths. Although I take high credence in truths to be an epistemic good which credence formation can promote (more shortly), I do not want to build a truth requirement into my notion of what it is for a proposition to be possible for an agent, because I want to leave open the possibility that an agent can be certain of a false proposition (as all too many agents, even ideally rational ones, presumably are). So I am sidestepping the debates about epistemic modals, epistemic relativism, and so forth, in order to concentrate on the purely doxastic aspects of rational credence formation.

[22] My account of the objects of credence as doxastic possibilities is bare-boned but sufficient for my purposes; for a more developed view along similar lines see Chalmers (2011).

[23] For discussion see Williamson (2000).

[24] Grasping may well have finer-grained variants. Perhaps an agent has a fine-grained version of grasping in the proposition P v Q if, say, she grasps P but not Q.

[25] This is to be distinguished from pragmatic reasoning, by which an agent attempts to achieve her ends.

conferred by features of the agent's mind which are not also features of her environment.[26] These features need not be things of which the agent is aware, or even of which she can in principle become aware. Internalist value is the sort of value which an agent's credences have regardless of whether the agent is in a skeptical scenario: if two agents, one of whom is in the good case and the other of whom is in the bad case, have identical mental states, then they are equal in the internalist epistemic value they possess. A form of epistemic value is *externalist* just in case it is not internalist: that is, just in case it is conferred at least in part by features independent of the agent's mind. Thus, of our two agents with identical mental states, only the one in the good case will have credences with externalist epistemic value.[27] The internalist/externalist distinction is traditionally important because it isolates a form of epistemic value, internalist value, which skeptical scenarios cannot impugn.[28]

Some of the epistemic goods discussed here are internalist and some externalist. But my interest in them has little to do with how they map onto the internalist/externalist distinction, which certainly does not exhaust the array of epistemic good-making properties. I focus instead on three others: rationality, certainty in truth, and propositional grasping. My thesis focuses on rationality, but it is important to begin by stepping back and showing how rationality is situated with respect to these other epistemic goods.

Though the internalist/externalist distinction enjoys little discussion here, there is another distinction which is important for characterizing the three sources of epistemic goodness I discuss. It is the distinction between individual credences and credence distributions as a whole. The epistemic goods I isolate may be predicated of either. When predicated of individual credences, they are *local* epistemic goods; when predicated of entire credence distributions, they are *global* epistemic goods. The following sections introduce the goods of rationality, certainty in truths, and propositional grasping.

[26] Thus if some mental states have environmental components, as Williamson (2000) thinks, the environmental components of these states have nothing to do with internalist epistemic goodness.

[27] Internalists include Steup (1999), Conee and Feldman (2001); Alston (1988); externalists include Goldman (2001), Williamson (2000), Bach (1985), Dretske (1970). I am not distinguishing between internalism/externalism about justification and about knowledge.

[28] Mainstream epistemology has a second internalism/externalism distinction, which circumscribes internalist value more narrowly and externalist value more broadly. We'll call an epistemic value *narrowly internalist* just in case it is conferred by features of which the agent is or can on reflection become aware, and *broadly externalist* just in case it is not narrowly internalist. This distinction is important for reasons not pertaining to skepticism. Narrow internalists include Chisholm (1977) and Ginet (1975). Of the epistemic goods we discuss here, however, none is strongly internalist. Instead, two (rationality and propositional grasping) are internalist in the broader sense, and one (certainty in truths) is externalist in the sense opposed to this. So I will leave narrow internalism and broad externalism to one side.

Rationality

Epistemic rationality (henceforth, just 'rationality') has a diffuse pedigree which spans a plurality of forms of justification. As such, rationality in some form is traditionally regarded as a component either of knowledge itself or of canonical instances of knowledge.[29] Knowledge, however, is a notion which is particularly resistant to importation into fine-grained epistemology, and I am not going to fabricate an analogue;[30] it is questionable whether fine-grained epistemology is even the worse without one, but I won't discuss the issue here.[31] Dissociation from knowledge by no means lessens the epistemic importance of rationality, at the same time that fine-grained epistemology is no poorer for lacking an analogue to knowledge. Rationality, I take it, is a sui generis epistemic good. Although it sometimes may aid agents in obtaining other epistemic goods (such as certainty in truths; see below), it is a source of goodness regardless of whether it ever does this. On some views, justification in coarse-grained epistemology is similar, being valuable independently of whether it conduces to or is a component of knowledge.

What, then, is rationality? It is the property which governs the formation and maintenance of doxastic attitudes, which normatively proceed by way of acquiring *evidence* (see Chapter 3). It is to be contrasted with pragmatic rationality, which governs the way in which an agent determines how to act in order to promote her ends. Sometimes an agent may have as an end the refinement of her credences to better match reality (on some sense of matching which I won't elucidate), but whether she is (epistemically) rational does not depend on whether she has such an end. Rationality as I understand it[32] is an internalist epistemic good, in the sense described above.[33] Rationality admits of degree, which there may be a variety of ways of measuring. But this thesis concerns itself only with the highest degree, ideal rationality.

[29] A non-canonical instance might be the true belief of a child, which some might think does not admit of rational evaluation at all (unless one agrees with me that children are epistemic agents).

[30] One might think that credence 1 in a true proposition is a suitable fine-grained analogue of knowledge. But this suggestion makes the dubiously strong requirement that an agent must be *certain* of a proposition she knows. Everyday cases of knowledge, by contrast, do seem to include situations in which the agent is less than certain of a proposition even though she knows it; take Dretske's (1970) agent who knows that the animal in the enclosure is a zebra even though she is not certain that it is not a painted donkey. Knowledge need not be certain.

[31] Williamson (2000) develops a knowledge-first epistemology with Bayesian characteristics, but the doxastic attitude with which he is concerned is coarse-grained; credences do not factor into his view.

[32] The claim that rationality is an internalist good is controversial. For externalist views specifically of rationality (as opposed to justification) see Plantinga (1993) and Brewer (1999).

[33] According to Bayesianism, as we will see, rationality is internalist but not narrowly internalist, since an agent can be rational in the Bayesian sense – by maintaining coherence and by conditionalizing – without being aware that she is.

Predications of rationality are often subject to a distinction between synchronic and diachronic.[34] *Synchronic rationality* is predicated of credences at a time. That is, it is intrinsic to a time, in the sense that all of its components are present at a single instant. *Diachronic rationality* is predicated of the process by which an agent manages (and perhaps changes) her credences from one time to another. This process is what I'll call an *update*; Chapter 2 gives more detail, but suffice it for now to note that an update need not be a *change* of credence, though it often is; it is, more generally, a re-assessment of one's credences in the light of new information. The components of diachronic rationality are spread out over multiple instants.

Does the synchronic/diachronic distinction provide a natural way to think about rationality? It would seem so. Synchronically, your high degree of confidence that the football match will proceed might be said to be rational (say, because you are certain that it will not rain and that rain is the only thing that might cause a match to be cancelled), or irrational (say, because you have a high degree of confidence that it will rain yet are certain that matches are always cancelled when it rains). Diachronically, the process by which you arrived at your degree of confidence might be said to be rational (say, because you formed a high degree of confidence that the match will proceed after hearing a reliable weather forecaster predict a 100% chance of sun), or irrational (say, because you formed a high degree of confidence that the match will proceed even though your evidence indicated that the match will not proceed). The philosophical notions of synchronic and diachronic rationality, though not distinguished in everyday language, do capture these intuitive distinctions.

Thus a complete theory of rationality will treat both synchronic and diachronic rationality. For reasons of space, however, mine focuses on diachronic rationality. Chapter 2 will give the Bayesian necessary condition for synchronic rationality, which will be enough to get my theory of diachronic rationality up and running. But I'll end this introductory section on rationality by stating a general principle which underlies the theory of diachronic rationality to be presented in subsequent chapters.

Evidentialism: A Constraint on Diachronic Rationality

The notion of evidence is central to a theory of diachronic rationality, but there is controversy about precisely how evidence and rationality relate. Pending Chapter 3's account of evidence, I can only speak in general terms. But I can give the following foundational Expansive Bayesian claim about diachronic rationality:

[34] For discussion see Christensen (2004).

> The receipt of new evidence (i) rationally mandates an update, and (ii) is the only thing which rationally permits one. (*Evidentialism*)[35]

An agent's 'new' evidence (see Chapter 3) is a proposition which is made salient to her by a representational experience, which is an experience as of things being some particular way. Evidentialism says that an agent may only update her credences when she receives new evidence. But it also says something stronger: that, on receipt of evidence, she is rationally obligated to update.

Why endorse Evidentialism? I endorse its mandating clause, (i), because I take it that evidence, in virtue of arising from a representational experience, is a sort of message from reality which it is irrational to ignore. The mandating clause is not as strict as it may seem, however. It does not force an agent to *change* her credences on receipt of new evidence; it says only that she must update them. An update is compatible with retaining the same credences after the update as she had before it. An agent may do this simply by assigning to her new evidence the same credence which she had before that evidence was made salient to her by the experience; this will, assuming that she conditionalizes (see chapter 2), issue in post-update credences which are the same as her pre-update credences. An update, then, is an assessment of one's present credences in the light of one's evidence, and this re-assessment often issues in numeric change, but need not; Chapter 2 says more about this.

What about Evidentialism's permissive clause, (ii), that nothing but evidence permits an update? I endorse this because evidence comes from a representational experience and is thus uniquely placed to affect the agent's representation of reality, which I am thinking of as the set of propositions in which she has a credence (her C-set). One might think that the notion of evidence can be bypassed and that the mere having of a representational experience suffices to permit rational updating. But evidence is an important intermediary between an agent's experience and her credence, for one might undergo a representational experience without having any doxastic attitudes pertaining to its veridicality. One may suspend judgment completely from any propositions which the experience offers up for one's consideration. These propositions themselves, by contrast, are candidates for truth or falsehood (and hence for the assignment of a credence), and as such, any credences which an agent assigns them are uniquely placed to influence her credences in the other propositions comprising her representation of reality.

This concludes my exposition of rationality. I have limited myself only to the most general comments in preparation for the specific account of rational updating to be given in subsequent chapters. It is a property which can be ascribed to the credences of epistemic agents, who are

[35] I am not using this term in the technical sense coined by Conee and Feldman (2004).

sentient and capable of entertaining the objects of their credences, and governs the ways in which they form and manage them. Before zeroing in on rationality and more specifically on diachronic rationality, let us situate rationality in the context of two other important epistemic goods.

Certainty in Truth

The second epistemic good is *certainty in truth*, where truth is predicated of propositions.[36] Certainty in truth is an externalist source of value, because whether an agent has it depends on features of the world outside of her mind, in particular whether the proposition in which she is certain obtains.

Whether a proposition is true or false makes a normative difference to what attitude one should have toward it, or so I assume. In coarse-grained epistemology, the epistemic goodness of truth is more straightforward than in fine-grained epistemology. The reason is that *beliefs* are isomorphic to propositions in a way which allows truth or falsehood to be predicated of them too. Credences, however, are not isomorphic to propositions: they admit of degree, whereas propositions do not. Thus credences themselves, unlike beliefs, cannot be true or false.[37][38]

For this reason, in fine-grained circles, truth has traditionally taken a position somewhat left of center.[39] Even so, there are several views about its role. Some argue that fine-grained epistemology has little to do with truth at all, and that Bayesianism is exclusively a mechanism for generating a fine-grained analogue of logical consistency.[40] Others argue instead that truth has a role to play as a long-term, though not short-term, goal of credence formation: that two agents who start off with wildly different prior credences will ultimately see their credence functions converge on the truth as they are presented with more and more (of the same) evidence, but this "washing out of priors" is acknowledged to be something that would never come close to occurring within the lifetime of human agents.[41] By whatever means these traditional approaches have tried to show fine-grained epistemology to be related (or unrelated) to truth, it is fair to say that rationality takes center stage.

However, recent work has tried to re-assert truth as a central epistemological value in fine-grained epistemology, by appeal to the notion of *truth utility*.[42] The idea is simple: the higher one's

[36] See Christensen (2004), Chapter 6, for an argument that high credence in truths is a distinct epistemic good from rationality.
[37] For a helpful discussion see Christensen (2004).
[38] Maybe there are some interesting things to be done with a non-standard semantics according to which truth values are degreed; I don't pursue this thought here.
[39] See e.g. Horwich (1992).
[40] See e.g. Howson (2000).
[41] For an insightful discussion see Earman (1992), Chapter 6. A related view, that coherence promotes 'reliability', is advocated by Olsson and Schubert (2007).
[42] See Joyce (1998, 2009), and Greaves and Wallace (2006). What I am calling 'truth utility', the literature tends to call 'epistemic utility'. To my mind this underpins a bias in favor of truth over and above the other two

credence is in a truth, the better (all else equal). And having more such valuable attitudes is, in fine- or coarse-grained epistemology, better than having fewer, as the treating of (coarse-grained) omniscience as a perfection attests. Truth-utility theory was originally developed as a means of justifying Bayesianism. But whether or not it does so convincingly, it is an effective way of spelling out the epistemic value of truth on a fine-grained framework.[43]

Recall that rationality has a synchronic specification and a diachronic one: credences at a time can be rational or irrational, and so can the process by which they move through time. Is the good of certainty in truths similar? In one way, it is: credences can have truth utility at a time, and they can maintain (or gain or lose) truth utility over time. But in another way, rationality differs from certainty in truth: there is comparatively little of interest to say about what it takes to *maintain* the maximal truth-utility through time, once one has it: one must merely remain certain of all of the truths. The diachronic interest in truth utility lies in figuring out how to accumulate more of it: how to move from an inferior state vis-à-vis truth to a superior one. In the case of rationality, by contrast, there are libraries, especially in formal epistemology, on how to maintain (ideal) rationality over time. This is the epistemology of *rational updating*, and it is to these libraries that this work will be a contribution.

Why is rational updating of such interest, especially in comparison with the relatively little interest shown in maintaining maximal truth utility? On the face of it, both seem arcane topics to be interested in at all, especially in a world where *no* agents have either maximal truth utility or (ideal) rationality (except perhaps in very localized areas of their credence distributions). One suggestion is that maintaining rationality over time requires doing something, whereas maintaining certainty in truths over time requires doing nothing (especially not forgetting). What must one do in order to maintain one's rationality over time? One must adjust one's credences in light of new evidence. Receipt of evidence, according to Evidentialism, makes a constant re-assessment of one's credences necessary, and one false move can demote one from (ideal) rationality. Maintaining rationality over time is about fielding goals that are under constant attack, whereas maintaining certainty in truth

---

epistemic goods I discuss, rationality and grasping, which I am discussing here. Hence my more egalitarian preference to use the term 'truth utility' instead.

[43] One might think that 'truth-utility functions', which quantify the epistemic value which accrues to credences to the extent that they come close to certainty in truths, are an aspect of *rationality*: that some such functions are more rational than others. This suggestion is developed, if not in precisely these terms, by contributions to the truth-utility literature since Joyce's (1998). That literature has since swung in the direction of what is now more broadly called *epistemic utility theory* (see e.g. Greaves and Wallace, 2006), where an *epistemic* utility function combines the epistemic values of truth *and* something along the lines of rationality or reasonableness into a single, all-epistemic-desiderata-considered utility function. This new style of epistemic utility theory depends heavily on arguments about what makes a utility function appropriately 'epistemic', and might provide a promising approach to developing a robust extension of the Bayesian notion of rationality. But epistemic utility theory, though compatible with what I will say, will not feature in the conception of rationality I develop here.

over time is about keeping up the status quo. This distinction may help explain why staying rational is of greater interest than staying certain in truths.

Let us finish our exploration of the normative epistemological domain, moving on to our third epistemic good which we will discuss more briefly, propositional grasping.

Propositional Grasping

Rationality and truth (or certainty therein) are bulwarks of traditional epistemology, in both its coarse- and fine-grained variants. But the focus on these two goods is lopsided. Propositional grasping is a third and very important epistemic good. I have already characterized it above, and will add only that it is an internalist value.

Propositional grasping has not received much publicity in either fine- or coarse-grained epistemology.[44] In the case of fine-grained epistemology, the reason seems to be that Bayesian models tend to work with idealized agents who are mere credence-forming machines. Most often the issue of propositional grasping is simply not raised. But when it is, it is assumed either that ideal Bayesian agents do grasp all propositions, or that credences are the kind of thing which can be had toward propositions one does not grasp.[45] Even so, some attention has been given to grasping in confirmation-theoretic contexts,[46] where cases arise in which the set of propositions over which an agent's credences are defined changes because the agent comes to grasp, and then assign a credence to, a proposition which she had not grasped previously, forcing her credence distribution to accommodate it somehow. But the main energies of Bayesian epistemologists have been expended in accounting for rationality in abstraction of grasping. The coming to grasp of new propositions, however, plays an important role in my Expansive Bayesian theory of rational updating. In particular, Chapter 3, Part III argues that coming to grasp a new proposition provides an agent with evidence which she could not have anticipated and hence opens up opportunities for forms of rational credence change which might be unexpected on a Bayesian theory of rationality; Chapters 5, 6, and 7 deploy this observation to respond to some criticisms of Bayesianism.

Of course, propositional grasping is necessary for the achieving of either of the other epistemic goods: one cannot have a credence in a proposition, let alone a rational credence or a certain credence in a true one, if one does not grasp that proposition (for an argument see Chapter 3). But

[44] For an exception see Williamson (2000), but his epistemology is at root coarse-grained.
[45] Van Fraassen (1989), chapter 13, discusses but does not endorse a view along these lines. Most often such views are not endorsed explicitly, though Eagle gestures at such a view (2011, 215).
[46] See Earman (1992) for a good discussion of this and the related issue of the 'problem of new theories'.

other things which are not epistemically interesting are necessary for being certain of truths too, such as breathing. Propositional grasping has epistemic value in its own right.

There is value in grasping a proposition even if one suspends judgment toward it, and even if one is already certain that it is false. There is also value in doing so even if one has an irrational credence in that proposition. The value lies in having fine-grained conceptual capacities, enabling one mentally to carve up modal space into very small pieces. The more propositions one grasps, the more variegated one's doxastic possibilities become, such that, even if one has no idea whether a given possibility obtains, or is even certain that it does not, one has a modal sort of epistemic success: One understands more precisely how things might or might not be. One is able cognitively to individuate a greater number of possibilities from each other. The value of grasping may sometimes of course be superceded by other forms of epistemic value, as when (for example) coming to grasp a new proposition requires so much intellectual effort that one fails to be rational or loses out on opportunities to raise one's credence in truths. I won't talk about the relative epistemic value of the three goods, except to say that there will be situations in which non-ideal agents may have to forego one for the sake of another.

Propositional grasping can also be thought of in synchronic or diachronic terms. I won't commit here to any views about the way in which propositional grasping confers epistemic goodness synchronically. There are surely at least as many of these as there are truth-utility functions. Concerning the diachronic goodness conferred by grasping, all I will claim is that one improves diachronically when one comes to grasp a new proposition. Thus all credences trivially have the synchronic good conferred by propositional grasping. It is epistemically good to change one's doxastic attitudes by coming to grasp a proposition which one did not grasp before.

Epistemically Ideal Agents

I have been saying that Expansive Bayesianism is a theory of ideal rationality, and hence will characterize the credence distribution of an ideally rational agent. But we have now identified two other epistemically ideal-making characteristics. For what comes later it will be important to show in the most general terms how they relate to rationality.

An agent is *rationally ideal* at a time just in case all of her credences are rational at that time.[47] An agent is ideal vis-à-vis *certainty in truths* at a time just in case all of his credences are either 1 or 0, such that the credences of 1 are in the truths and the credences of 0 in the falsehoods. An agent is

[47] Christensen (2007) has isolated a problem with the notion of ideally rational agents, but this problem arises only when they are uncertain whether they are ideally rational, and it is not at all clear that ideal rationality is compatible with lack of certainty in this regard; I won't discuss the issue here.

an *ideal grasper* just in he grasps all propositions. Thus being an ideal grasper is defined relative to all propositions, whereas being ideal in either of the other two respects is defined relative to an agent's credence distribution.

What are the logical relations among these ideals? First, these ideals are compatible: the claim that an agent is ideal in one respect is consistent with the claim that he is ideal in any of the others; moreover, an agent may possess all of the ideal-making characteristics simultaneously.[48] Second, that an agent is ideal in one way does not entail that she is ideal in any of the others. An agent may be a certainty-in-truths ideal without being a rational ideal: her credences in the truths can all be 1 and in the falsehoods can all be 0 and yet she can be irrational if this state is not an appropriate response to her evidence.[49] An agent can be an ideal grasper without having any credences at all: she may simply entertain every single proposition without having any attitudes expressing how likely it is to be true; in this case neither ideal rationality nor the truth ideal are options for her, since these both require having some credences.[50] Finally, an agent may be rationally ideal without having all of her certainties be in truths or in the negation of falsehoods. She may also be rationally ideal without grasping all propositions.[51]

The latter claim will be important in what follows, so I will name it, as the negation of the claim that

Being ideally rational entails grasping all propositions (*Rational Grasping*)

Not-Rational Grasping has some far reaching epistemological implications which Chapter 3, Part I, spells out. It will be the first step in building a theory of rationality for sentient agents who mentally engage with the objects of their credence rather than mechanically crunch numeric credal data.

The final section pictorially represents the logical relations among these epistemic ideals.

---

[48] Limited agents may sacrifice proximity to one ideal for proximity to another. A move toward certainty in a truth may demand moving away from rational credences (say, by a lucky but irrational guess), and a move toward having rational credences may result in a move toward falsehood (say, in skeptical scenario). And a move toward grasping might come at the expense of either rationality or truth, since coming to grasp new propositions costs time that may otherwise be spent in trying to make the credences one already has more rational, or by otherwise trying to move toward certainty in truths. Similarly, the efforts expended in doing the latter things may limit one's ability to grow in grasping.

[49] If being rationally ideal, as some extreme Bayesians think, is just being probabilistically coherent, then, since an agent with a credence of 1 in all truths and of 0 in all falsehoods is probabilistically coherent, she is also rationally ideal. This is not my view: I think there is more to rationality at a time than coherence.

[50] Since a way in which things could be is a doxastic possibility, this raises the question of the metaphysical status of doxastic possibilities which are not entertained by an agent; it is akin to the status, whatever it is, of experiences which are not had by an agent. Perhaps they are experiences which she has in other (logically) possible scenarios.

[51] She need only grasp those over which her credences are defined, which, given our conception of the objects of credence, she does trivially.

Epistemic Properties of Propositions

For what comes later it will be important to distinguish different sets of propositions over which an agent's credences may be defined, and to do so in terms of the epistemic properties which these propositions may have. There are eight sets of propositions which are relevant to distinguish:

(1) the set P of all propositions *simpliciter*;

(2) the set C of propositions in which the agent has some credence, no matter what its value. C is what we'll call *the set over which the agent's credence function is defined*; it is equivalent to what I called, above, the agent's 'representation of reality';

(3) the set E of propositions in which the agent has an extremal credence, that is, a credence of 1 or 0.

E is divisible into two exclusive and exhaustive subsets:

(4) one subset of E of propositions in which the agent's credence is 1 (*the 1 subset*), and

(5) another in which his credence is 0 (*the 0 subset*).

The 1 and 0 subsets are themselves each divisible into two mutually exclusive and exhaustive subsets:

(6) the intersection of the 1 set with the set, T, of all truths (comprising exactly half of the propositions in P);

(7) the intersection of the 1 set with the set, F, of all falsehoods (comprising exactly half of the propositions in P);

(8) the intersection of the 0 set with T; and

(9) the intersection of the 0 set with F.

Intersecting all of the other sets is:

(10) the set G of propositions which a given agent grasps.

Finally, intersecting all of other sets is:

(11) the set R of all of the propositions in which an agent's credence is rational.

We may visualize this setup as a Venn diagram consisting of a series of concentric circles (see Fig. 1).

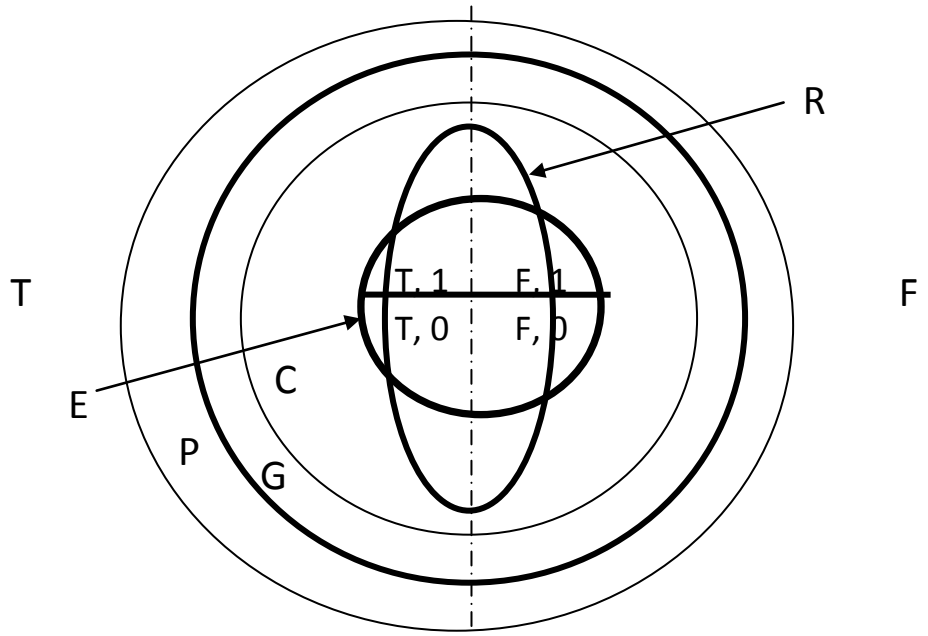

Fig. 1. The sets of propositions which are possible objects of credence.[52]
P: the set of all propositions *simpliciter*;
T: the subset, comprising exactly half of P, of all truths;
F: the subset, comprising exactly the other half of P, of all falsehoods;
C: the set of propositions over which the agent's credence function is defined
E: the set of propositions in which the agent has an extremal credence;
T, 1: the intersection of the 1.0 set with T, the set of all truths (which is a proper subset of P consisting of half of the elements of P); this set contains all of the truths of which the agent is certain;
F, 1: the intersection of the 1.0 set with F, the set of all falsehoods (which is a proper subset of P consisting of the complement of T); this set contains all of the falsehoods of which the agent is certain;
T, 0: the intersection of the 0 set with T; this set contains all of the truths which the agent is certain are false;
F, 0: the intersection of the 0 set with the set F; this set contains all of the falsehoods which the agent is certain are false;
G: the set of propositions which the agent grasps.

[52] The representation is overly neat; a representation of the possible objects of real credences would be much less schematic; in particular, G might comprise the set-theoretic union of disjoint sets dotted all over the diagram.

R: the set of propositions in which the agent has a credence which is proportioned to her evidence, i.e., which achieves the maximal proportioning utility. (It is for purposes of visual representation only that R must occupy two disjoint areas in the diagram.)

The largest set, P, is divided into two exclusive and exhaustive subsets, the subset T (on the left) and the subset F (on the right). C is a proper subset of P, and E a proper subset of C. E is quartered into four exclusive and exhaustive subsets: the intersection of T and the 1 subset, the intersection of T and the 0 subset, the intersection of F and the 1 subset, and the intersection of F and the 0 subset. G intersects all of the other sets but neither contains, nor is contained by, nor is equivalent to, any of them. R intersects all of the other sets except for P, but neither contains, is contained by, nor is equivalent to, any of them.

Figure 1 depicts an epistemically imperfect agent at a single time who possesses none of the ideal-making features. This depicts the fact that the three epistemic ideal-making characteristics can come apart, as we have said.

Conclusion

This completes my exposition of the terms of my discussion and of the main general features I take the fine-grained epistemological domain to have. I have characterized the notions of epistemic agent, credence, object of credence, epistemic good, and epistemic ideal, as well as three salient epistemic goods which, when possessed perfectly, are also epistemic ideal-making features. We are finally in a position to zero in more closely on ideal rationality.

2. Introducing Bayesianism

Introduction

Bayesianism has been regarded alternatively as a descriptive framework for how scientists typically reason (Howson and Urbach, 1993), and as a normative theory, about how scientists (or anyone) ought to reason. As should be clear by now, I am using it to build a normative theory, which tells us what ideal rationality is like.

Bayesianism gives a necessary condition for synchronic rationality, and another for diachronic rationality. The synchronic condition is:

> An agent is rational at a time only if his credences are a probabilistically coherent function. (*Bayesian Statics*)

The diachronic condition is:

> An agent is rational from one time to another only if, when she updates on evidence, she does so by conditionalizing. (*Bayesian Updating*)

The view I defend, Expansive Bayesianism, endorses Bayesian Updating and a slightly weakened version of Bayesian Statics:

> An agent is rational at a time only if his credences are a subset of a probability function which is defined over all of the truth functional combinations of the propositions in which he has credences. (*Weak Bayesian Statics*)

This chapter clarifies all three claims, in preparation for the development of Expansive Bayesianism.

Bayesian Rationality at a Time

I'll present Bayesianism's condition on rationality at a time, and then motivate Expansive Bayesianism's weakening of it.

> An agent is rational at a time only if his credences are a probabilistically coherent function. (*Bayesian Statics*)

The credences which the agent has at any given time are defined over a set of propositions, which the previous chapter called the agent's 'C-set' (or just named 'C'), short for 'the set of propositions in which the agent has a credence'.

Bayesian Statics says that the numeric values of the agent's credences must be derivable by a *function* from the propositions in C – in particular by a *probability* function. Thus the Bayesian constraints on synchronic credence are none other than the axioms of the probability calculus, and the rational agent's credences in the propositions in C are a probability distribution over C. The probability values, however, are not the chances or frequencies of physics or statistics. Instead, they are expressions of an agent's degrees of confidence in the propositions over which they range. Hence another name for credence is *subjective probability*: if I have a credence of *x* in *p*, I have a subjective probability assignment of *x* that *p* obtains. That assignment is determined by my *credence function* (*Cr*), which takes a proposition as input and gives a unique credence as output.

Bayesian Statics holds that Cr must be *probabilistically coherent*, which means that it must obey the axioms of the probability calculus. These are:[53]

> For all propositions *p* which are elements of C, $Cr(p) \geq 0$. (*Non-negativity*)
>
> If *p* is a tautology and is an element of C, then $Cr(p) = 1$. (*Normality*)

[53] Eagle (2011).

For any countable sequence of propositions p_1, …, p_n, …, such that any two propositions in the sequence are mutually exclusive, and each p_i is an element of C,

$$Cr\left(\bigvee_i p_i\right) = \sum_i Cr(p_i)$$

(*Countable Additivity*)

Let's look briefly at each axiom.

Non-negativity. Non-negativity establishes a credence of 0 in any given proposition *p* as the most certain an agent can possibly be that not-*p*.

Normality. Since a credence carves the agent's total doxastic space into proportions, measuring credences requires a *normed measure*, i.e., a scale which has a standard maximum that represents the total possible value of which each credence is a proportion (an example of a measure which is *not* normed is weight). The Normality axiom establishes 1 as this maximum. Thus together with the other axioms, Normality establishes a credence of 1 is the most confident that an agent can be that P.[54] Since the tautology obtains in every possible world and since this can be seen merely by understanding the logical connectives in it, no doxastic possibility space should exclude it and thus the agent should thus have credence 1 in it.[55] Why 1 and not, say, 2 or 1.7? There is nothing special about the unit interval. It is just the easily divisible number chosen as a handy representation of the array of credal attitudes. Thus having a credence of, say, 1.7 is not a way of being probabilistically incoherent; it just fails to possibly represent a credence at all on this framework.

Countable Additivity. The final probability axiom, Countable Additivity, establishes a probabilistic version of consistency. According to Normality, the total credences which a rational agent may spread over a proposition and its negation (or, more generally, over a *logical partition*, i.e., a mutually exclusive and jointly exhaustive set of propositions) cannot exceed 1. Countable Additivity adds that they cannot be less than 1, either.[56]

---

[54] Proof: Given the Normality axiom, one's credence in *P or not-P* has to be 1. Given the Countable Additivity axiom, one's credence in *P or not-P* has to equal one's credence in *P* plus one's credence in *not-P*. Given the Non-negativity axiom, the minimum value for either *P* or *not-P* is 0. If you're adding two non-negative numbers together to get 1, the biggest either of them can be is 1.

[55] This suggestion might seem surprising given that the objects of credence are as fine-grained as Chapter 1 claims them to be, on the grounds that a proposition might be doxastically possible which is logically or metaphysically impossible. I'll address this worry at the end of this section, after I've finished introducing Bayesian Statics.

[56] One might reject Countable Additivity, in particular in favor of Additivity (which holds for finite sets of propositions), the axiom that credences must sum to 1 over a finite set of propositions, is that doing so simplifies the mathematics significantly. Haenni (2009) contests Countable Additivity on the grounds that it does not appropriately model ignorance. Some forms of ignorance may not be ideally rational, in which case one would not expect a theory of ideal rationality to model them. But some are compatible with ideal rationality, a case which Chapter 1 has begun to make; one example is failing to grasp a proposition, and

Note that that I have stated each axiom with a clause relativizing it to a limited domain of propositions, namely, the agent's C-set. It is common for Bayesian accounts to state the axioms over an unrelativized domain, as governing the agent's credences in *all propositions* full stop.[57] The latter way of doing things, however, entails that

Rational agents have credences defined over all propositions. (*Complete*)

Complete, in turn, carries some untoward consequences about the nature of credences and of epistemic agents. In particular it entails that the rational agent has credences in all propositions. If credences are in grasped propositions only (which Chapter 3, Part I argues), then this entails that the rational agent *grasps* all propositions. Chapter 3, Part I argues that Complete is false. Moreover, Chapter 1 has argued that grasping all propositions (or: meeting what I called the 'grasping ideal') is a form of epistemic perfection which seems to have little to do with rationality and should not be built into an account of it.

The next section motivates a slight weakening of Bayesian Statics as the first component of the theory of rationality I'll endorse, Expansive Bayesianism.

Weak Coherence

Even relativizing the probability axioms to a domain yields an idealization too many. For probabilities are defined over sets of propositions[58] which have particular properties. One such property is that of being truth-functionally complete: they are defined over all of the truth-functional combinations of their atomic elements. Given Countable Additivity, this means that Bayesian Statics requires even an agent with credences in a small number of atomic propositions to have credences defined over propositions which are infinitely complex. But rationality, Chapter 1 suggested, does not mandate *having* particular credences; instead, it governs how one handles and manages the credences which one does have. Stipulating what credences agents ought to have, then, lies outside the bounds of rationality strictly speaking – hence Bayesian Statics, which entails that agents have credences in all of the truth-functional combinations of the atomic propositions in which they have credences – goes beyond the remit of rationality.

---

another is grasping a proposition yet assigning no credence to it on account of being completely in the dark about how likely it is to be true.

[57] Eagle (2011) and Howson and Urbach (1993) however also explicitly restrict the domain of the probability function.

[58] Probabilities *qua* credences (as opposed to, say, physical chances) can of course also be defined over sentences but I am defining them over propositions instead, because I think that making the domain language dependent is unnecessarily limiting; Chapter 3, part III provides some reasons.

What would be objectionable about construing rationality in terms of Bayesian Statics? Chapter 3 discusses this issue at greater length, so the present introduction will be cursory. The problem arises when we consider the claim:

> Having a credence in *p* entails grasping *p*. (*Grasping*)

Either Grasping is true or it is not. If it is true, then Bayesian Statics gives us as a rational ideal an agent who is so idealized that she grasps even infinitely complex combinations of propositions. But grasping is not a feature of rationality either, as Chapter 1 argued (although it is an epistemic good in its own right). Moreover, subsequent chapters will show that some of the criticisms launched against Bayesianism (Chapters 6 and 7) can be diffused if we take on board that ideally rational agents can come to grasp propositions which they do not already grasp (including the truth-functional combinations of atomic propositions which they grasp already). But if Grasping is false, then Bayesian Statics gives us as a rational ideal an agent who is so mechanical as to have doxastic attitudes without a mental life to accompany them. Credence is severed from the mind, at which point it becomes hard to imagine what makes it a *doxastic* attitude at all. Whether Grasping is true or false, the picture of ideal rationality resulting from Bayesian Statics extends beyond the remit of rationality strictly speaking.

There is a simple solution. Instead of endorsing Bayesian Statics as it stands, with its commitment to a full-fledged algebra of propositions, I'll endorse

> An agent is rational at a time only if his credences are a subset of a probabilistically coherent function formed by defining probability values for all of the truth-functional combinations of the propositions in which he has credences. (*Weak Bayesian Statics*)

Let's call a credence distribution which obeys Weak Bayesian Statics, but not Bayesian Statics, *weakly coherent*. Weak Bayesian Statics diverts the dilemma that arises over Grasping. It yields a rationally ideal agent who grasps all of the propositions over which his credences are defined. It requires neither that rationally ideal agents grasp infinitely complex propositions, nor that there be doxastic attitudes toward objects which agents do not grasp. The reason is, it does not require that ideal agents' credence functions be bona fide probability functions; it requires only that probability functions be recoverable from them.[59]

[59] Of course, infinitely many probability functions can be recovered from a credence distribution over a domain which is not an algebra. But since they will not in fact be the credences of the agent in question, this is not a problem.

Weak Coherence: Necessary but not Sufficient

Weak Bayesian Statics gives a necessary condition on credence at a time. Many also think that it (or Bayesian Statics) gives a sufficient condition.[60] This view is extreme. It yields the result that *any* degree of confidence in any proposition, no matter how ludicrous it strikes us as being, is rational as long as the agent's credences obey the probability axioms. For this reason I regard the synchronic constraint on rationality as necessary but not sufficient.

Some would add the condition that an agent's credences at a time be *reasonable*. 'Reasonable' can be spelled out in different ways, say in terms of particular alleged confirmational virtues such as simplicity or explanatory power[61], as corresponding to any objective chances there may be,[62] as being maximally non-committal[63], in terms of manifesting epistemic virtues,[64] or as unanalyzable[65]; but the basic intuition is that there is some class of coherent credence functions which is epistemically better in some respect than others. Those who endorse a reasonableness condition will note that many of the successes which Bayesian epistemology has enjoyed as a framework for confirmation theory depend on positing particular prior credences as the 'correct' ones.[66] My theory, Expansive Bayesianism, is neutral about what additional sorts of reasonableness constraints there may be, since its focus is on rational change of credence over time, but in endorsing only a necessary condition for synchronic rationality, Weak Bayesian Statics, it is open to some such constraint.

Before moving on to diachronic rationality, let's address a few objections to my account of synchronic rationality.

Objection from Metaphysical Necessity

One might suspect that there is a conflict between my account of the objects of credence and my use of the probability axioms to model rational credence at a time. The worry begins by recalling (from Chapter 1) that Expansive Bayesianism takes the objects of credence to be very fine-grained, so that an agent can have a different credence in *Hesperus is Hesperus* than she has in *Hesperus is Phosphorus*, even when 'Hesperus' and 'Phosphorus' rigidly designate the same item. This is a problem, alleges the objector, because Bayesian Statics requires that the agent assign the same

[60] See e.g. Jeffrey (1983 and 2004) and van Fraassen (1984 and 1989).
[61] Psillos (2007) suggests that explanatoriness is embedded in rational conditional credences.
[62] Lewis (1980).
[63] For discussion see Shimony (1985).
[64] Ramsey (1931).
[65] Williamson (2000).
[66] For discussion see Eagle (2011), p. 212, who however does not go so far as to say that there are any necessary reasonableness constraints on synchronic credence.

credence to each of these propositions. It would seem, then, that Bayesian Statics and the Expansive Bayesian account of the objects of credence could yield contradictory verdicts about the rational credence in such propositions.

But this worry incorporates an assumption which I do not share: that the tautology includes all *metaphysically* necessary truths, which is to say, propositions which are true in every possible scenario but whose negation is not a logical contradiction. *Hesperus is Phosphorus* is an example. Bayesian Statics, by contrast, understands the tautology as including only *logically* necessary truths, which is to say, propositions whose negation is a logical contradiction. *Hesperus is Phosphorus*, although false in every possible scenario due to the metaphysical nature of Hesperus and of Phosphorus (i.e. that they are identical), is not logically necessary: its negation is not false in virtue of the meanings of any of its truth-functional connectives. Agents may rationally have some credence in *Hesperus is Hesperus* – which must be 1 since its negation is a contradiction – and another in *Hesperus is Phosphorus* – which need not be 1 since it is logically possible that Hesperus is not Phosphorus, even if it is not metaphysically possible.

But why, the objector may persist, endorse a rational distinction between logically and metaphysically necessary propositions? Why does the tautology deserve no less than credence 1, whereas metaphysically necessary propositions can receive any credence at all? One reason is that a theory of rationality should be able to model scenarios in which an agent who, because he is unaware that (say) Hesperus is Phosphorus, has a credence of less than 1 in *Hesperus is Phosphorus*. Such situations can be rational, because the agent's uncertain credence is due merely to a lack of evidence indicating Hesperus's identity with Phosphorus, and not to any failure of the agent's ability to reason. A theory of rationality which demanded that agents have particular credences in a posteriori propositions, even metaphysically necessary ones, oversteps its bounds. It mistakes the goodness of certainty in truths (here, about metaphysical necessities) for the goodness of rationality, which Chapter 1 argued are distinct. There is thus no failure of rationality in investing a credence of less than 1 in a metaphysically (but not logically) necessary proposition.

Objection from Rational Uncertainty in Tautologies

But the objector might still claim that my view has trouble with tautologies, even construed as propositions of logical and not metaphysical necessity. For it can surely be doxastically possible for an agent that a tautological proposition be false, entailing that she assign it a credence of less than 1, and surely it can be rational to be uncertain about the status of a tautology. There are two ways in which this uncertainty might come about. First, the tautology may be too complex to parse easily.

Imagine a philosophical layman who, in contemplating a tautology such as ((((*P* → *P*) → *P*) → *P*) → ((*P* → *P*) → *P*) → *P*), is unsure about its status and hence assigns it a credence of less than 1. Expansive Bayesianism (and Bayesianism more generally), the thought is, says that everyday folk of this sort are irrational for questioning the status of hard-to-parse tautological propositions; but surely such a conception of rationality is unreasonably demanding. Another reason why one might be uncertain of a tautology is if one questions the status of classical logic itself. Take for example the non-classical logician who does not assign credence 1 to *P* v *not-P*, since she takes there to be a third alternative for which some credence must be left over; take the paraconsistent logician who does not assign credence 0 to *P & not-P*, since on her view it might be true. Expansive Bayesianism says that such philosophers are irrational – but surely, says the objector, supporters of classical logic cannot win that debate so easily! Thus there are, it would seem, situations in which it can be rational to assign a credence of less than 1 to a tautology, and my account of the objects of credence cannot accommodate them.

What is the Expansive Bayesian to respond? One might opt for an extreme response, arguing that it is never rational to be uncertain of a tautology, or that one cannot grasp a tautology without assigning credence 1 to it and hence the agents in the examples cannot have credences in the tautologies in question to begin with. But a more concessionary response, which I will opt for, admits that this objection zeroes in on a potential limitation of the Bayesian framework to model rational credence. Thus although Expansive Bayesianism construes propositions in such a way that an agent can have a less-than-certain credence in a tautology (since it can be doxastically possible for her that some tautology be false), Expansive Bayesianism is not in a position to model such credences as rational. However, any damage caused by this limitation is localized to specific debates about the status of tautologies or of classical logic itself. Bayesianism does allow rational credences in all kinds of other truths to be modelled, since, as we have seen, it does not demand that agents be certain of truths that are (putatively) metaphysically but not logically necessary, such as truths of mathematics, mereology, and fundamental ontology. This flexibility allows Bayesianism to model disagreement about anything other than classical logic – which is to say, about very much indeed.

This limitation is mitigated by the wide-lense observation that any epistemology modelled after classical logic will struggle to deal with views questioning classical logic itself. If Bayesianism is limited, so is every other formal epistemology which builds in classical logic. Moreover, it is worth pointing out that at this stage there is no well developed way to model credence in non-classical logic at all, let alone uncertain credence in tautologies. Bayesianism is a positive account of a way to formally model credences which are distributed over a classical representation of logical space; in order to substantiate an objection against Bayesianism on the grounds that it cannot model

uncertain credence in tautologies, a non-classical logician must produce an alternative way to model credences distributed over a non-classical representation of logical space. As yet there is no such account forthcoming: Bayesianism is as good as it gets.[67]

Nonetheless, these questions remind us that Bayesian epistemology is still a work in progress. My theory is Bayesian because of Bayesianism's great promise, not its perfection. One of my hopes is that employing a Bayesian framework for my own theory will help to smooth over at least some of Bayesianism's rough-and-ready patches in ways that might point the way toward addressing others as well.

Having sketched the Bayesian condition on synchronic rationality, and the Expansive Bayesian weakening of it, Weak Bayesian Statics, let us present the Bayesian condition on rationality over time.

Bayesian Rationality from One Time to Another

Recall the Bayesian necessary condition on rationality over time:

> An agent is rational from one time to another only if, when she updates her credences, she does so by conditionalizing. (*Bayesian Updating*)

In Bayesian circles, an agent's credences are sometimes called his *priors*, short for *prior credences*. This term has diachronic connotations: prior credences take on the value they do only prior to something's taking place. That something is an update on evidence. Bayesian Updating is a claim about how the agent's post-update credences derive rationally from his pre-update credences.

Update denotes the way in which an agent aligns her credences to accommodate new information. Any new information which warrants an update I am calling *evidence*. I am using 'evidence' as a term of art designed for fine-grained epistemology. My use of this term does not perfectly match the ordinary English usage, or even the way it is used in coarse-grained epistemology (which is arguably more germane to ordinary English than fine-grained epistemology). It is an apt term of art, however, because it plays a role which is a fine-grained analogue of the role which 'evidence' plays in English, at least when English is used to talk about what rationally legitimates or mandates a change of opinion. Evidence retains some features of coarse-grained philosophical notions of evidence, one being that it is propositional.[68]

[67] But see Robbie Williams (forthcoming) for a discussion of generalized probabilism in non-classical frameworks.

[68] For strong argumentation to this effect see Williamson (2000), Chapter 9, and Dougherty (2011), Chapter 14.

We may distinguish two varieties of evidence; whether a proposition is one or the other for an agent is time-dependent. An agent's *new evidence* at a time *t* is any proposition in which, as the result of what we'll call below an 'evidential experience', she forms a new credence. An agent's *background evidence* at *t* consists in any propositions in which her credence has remained constant in the face of the most recent evidential experience she has had before *t*.

A point about updating must be emphasized which will play an important role in the following pages. It must be emphasized because it tends not to be given its due weight in Bayesian literature, and hence, as we'll see later (in particular, Chapters 5, 6, and 7), results in the building of some implicit yet false and hobbling assumptions around the Bayesian framework. The important point is that *updating has two conceptually distinct components,* not one. One is the *input component*, in which the agent assigns a credence to her evidence; the other is the *adjustment component*, in which she adjusts her other credences in order to maintain probabilistic coherence. This distinction between input and adjustment is conceptual: we can conceive of a scenario in which an agent changes her credence in her evidence proposition without changing any other credences, and vice versa.

Bayesian Updating pertains to the adjustment step only, saying that it must proceed by conditionalization. It says nothing at all about the input step. That is, it says nothing about where the evidential proposition comes from or about what if anything regulates the credence which the agent assigns to it. Giving an account of the input step is thus one important way in which a Bayesian theory of rationality must build on the framework. The main concern of this dissertation is to argue in favor of a Bayesian account of the adjustment step. But a crucial part of my argument involves laying bare and rejecting some implicit assumptions about input. Hence the next section gives an account of how this step works, and Chapter 3 (Part II) discusses rational constraints on it. The final section of this chapter outlines the Bayesian take on the adjustment step, to prepare for the discussion of conditionalization in Chapters 4-7.

The Input Step

In the input step, the agent assigns a credence to her evidence proposition. Where does this proposition come from, and what determines the value which she assigns to it?

The input step, when rational, is always brought about by a *representational experience*, which Chapter 3 argues. Representational experiences represent or portray things as being thus-and-so:

they depict, describe, or are more generally "about", what is or could be the case.[69] A representational experience will make at least one proposition salient to an agent, presenting it as a candidate for being the case. Any such proposition becomes the agent's new evidence E. If the agent is rational (given Evidentialism, Chapter 1) she initiates an update by assigning a credence to E. What is important is that *the input step proceeds as the causal result of an experience*, as we'll see. This causal process is not an update itself; rather, it serves as the input for an update.

The agent may be in one of three mutually exclusive and exhaustive states immediately before performing the input step. (i) She has a credence in her evidence proposition E which is different from the credence she assigns in the input step; (ii) she has a credence in E which is the same as the credence she assigns in the inputs step; (iii) she has no credence in E at all.

Here is an example of case (i), which will be the most familiar to us, in which the input step is a change of a pre-existing credence: Yesterday, I walked outside in the direction of the rosebush, having a very low credence in *The rosebush has a third shoot*, since, when I saw it the day before, it only had two. But then, in noticing a third shoot, I had an experience as of its having a third shoot, and thus in the input step changed my credence in *The rosebush has a third shoot* – in particular, by raising it significantly. Cases of type (i) are canonical in the Bayesian literature. But for the purposes of my project it is important to distinguish two other types of case.

In case (ii), the agent has a prior credence in her evidence but does not change it when she assigns an input credence. For example, imagine that, today, I again walk toward the rosebush, seeing which prompts an identical experience as of its having a third shoot (and I've had no evidentially relevant experiences since yesterday); on being presented today with the proposition *The rosebud has a third shoot*, I assign as an input value the *same* credence which I already had; namely, the one I assigned yesterday on seeing the shoot for the first time. The experience has made the evidential proposition *The rosebud has a third shoot* salient to me again, but not in a way which motivates me to change the credence in it which I already had. Thus in case (ii), the receipt of evidence will not result in a *change* of credence, and hence neither will the update as a whole. Return to our example. Yesterday, when I noticed the shoot for the first time, I conditionalized on *The rosebush has a third shoot* and wound up with a higher posterior credence in *The rosebush is healthy* than before I noticed the shoot. Today, when I walk outside and *The rosebush has a third shoot* is presented to me again, I form the same credence in it as yesterday, once more conditionalizing *The rosebush is healthy* on it. In the example, my credence in neither proposition

[69] My account is neutral on the issue of whether experiences are intrinsically representational, in the sense of having veridicality conditions, though I am sympathetic to the claim that they do.

has changed, but I have still updated, because I have still assimilated information which is new in the sense of being called once more to my attention.[70]

Finally, in case (iii), the agent does not have a credence in E when she has the evidential experience: I walk outside suspending judgment about whether the rosebush has a third shoot, having no credence at all in *The rosebush has a third shoot ….* . When I see the bush and assign a high credence to *The rosebush ….* , it is for the very first time: I have assigned this proposition an initial credence from scratch. There are thus three mutually exclusive and exhaustive states in which an agent might be before the input step. They are important to distinguish for reasons that will become clear later, when we see how the adjustment step relates to the input step.

Examination of the input step might surprise those who think of Bayesian epistemology as rigid and prescriptive. For in the input step, the agent moves from one credence in an evidential proposition to another, just by having an experience. There is no conditionalization involved at all, or any prescriptive formula by which the assignment of credence must proceed. For all that the probability calculus has to say, the agent apparently gets her input credence for free. But if this is so, why the fuss over conditionalization as the only way in which the adjustment step may rationally proceed? It hardly seems to make sense to strictly regulate the adjustment step while remaining hands-off regarding the input step.

The answer to this worry is that, no indeed, Bayesianism does not regulate the input step. This is one reason why Bayesianism is a framework rather than a theory. It says that adjustment must proceed by conditionalization, but says nothing about the input which must be conditionalized on. A major contention of this work is that, although Bayesian theories have not explicitly accounted for the input step, many have built in implicit assumptions about it. Moreover, since input is conceptually prior to adjustment, implicit assumptions about input have far-reaching implications for a theory of updating as a whole. Chapters 5, 6, and 7 will show the detrimental effects which assumptions about input have had on Bayesian theories of updating, and will present my Expansive Bayesian theory of input and adjustment as an alternative. Chapter 3, Part II discusses a few rational constraints on the input step, proposing a few of its own and rejecting a few which are endorsed in some Bayesian circles.

The Bayesian framework is compatible with different kinds of constraint on the input step, as well as with no constraints at all. What *defines* Bayesianism, by contrast, is its take on the second step in updating, the adjustment step, in which the agent assimilates her other credences to accommodate any new credence she may have assigned to her evidence proposition in the input step. Bayesianism

[70] My credence in *The rosebush is healthy* might today be raised or lowered for other reasons, of course.

says that the adjustment step must rationally proceed by a process called 'conditionalization', and the rest of this chapter discusses what this amounts to.

The Adjustment Step: Conditionalization

To introduce the formula which governs conditionalization, it will help to recall the notion of conditional credence. This notion suggests a natural way to specify how one credence function should give rise to another when the agent receives new evidence. Conditional credence is defined as follows:

$$Cr(P \mid E) = \frac{Cr(P \,\&\, E)}{Cr(E)}$$

(*Conditional Credence, Definition*)

That is, the agent's credence in P conditional on E is given by her credence in *P & E* divided by her credence in E. Since an agent's conditional credence gives a value for her credence in P given E, a natural assumption is that, if the agent becomes certain of E, and if he has a credence in P conditional on E, then he should adjust his credence in P so as to equal his credence in P conditional on E. He should do something similar (but slightly more complex) if, instead of becoming certain of E, his credence in it changes just a little. This mechanism for changing one's credence function on receipt of evidence is called 'conditionalization',[71] and it is the *sine qua non* of Bayesian constraints on updating.

Bayesian Updating is a claim about how the (weakly) coherent agent updates his credences on acquiring evidence *e* (in which he has a new credence of *x*)[72]. Any such update must proceed by conditionalization. There are two formulas governing conditionalization (Bayes's Theorem and Probability Kinematics) which I'll discuss in a moment. First, a description in words to motivate them: If an agent has credences conditional on a proposition E which have different values than they do unconditionally, then those credences' values will change when he conditionalizes them on E; if he does not have credences conditional on E which differ in this way, then conditionalization will result in no change in any of the agent's credences; in this case the net change as a result of acquiring E as evidence will be the new credence in E (and of course, given Countable Additivity, the corresponding change in the agent's credence in not-E). Bayesian Updating logically presupposes at least Weak

[71] Sometimes 'conditionalization' is reserved for the case in which the agent is certain of her evidence. But as we'll shortly see, I use it more broadly to refer also to cases in which she is not.
[72] Henceforth, for ease of expression, I will take 'with new credence *x*' to be assumed when I talk about acquiring evidence.

Bayesian Statics, since conditionalization can only be performed by agents who are at least weakly coherent when they receive evidence. Thus it ranges only over agents who are already synchronically (ideally) rational.

'Conditionalization', as referred to by Bayesian Updating, is ambiguous between two different procedures, *straight conditionalization* and *Jeffrey conditionalization*. For ease of expression I use the generic term 'conditionalization' to refer to either, since it will always be clear which procedure the agent must follow. Whether the agent performs straight or Jeffrey conditionalization depends entirely on the value of his new credence in his evidence proposition. An agent performs straight conditionalization when his new credence in his evidence proposition E is 1. He performs Jeffrey conditionalization when his new credence in E is some value between zero and 1 exclusive (and when, as we'll see, he also has a credence in not-E or in each of a set of propositions equivalent to it).[73]

Straight conditionalization works like this. Since the agent has become newly certain of E, he must change his credence in any other proposition P so as to equal the value of his conditional credence ($P \mid E$), if he has one. A rough and ready definition of 'conditionalizing update', then, is this:

> A *conditionalizing update* is one in which the agent sets her new credences equal to the value of her old credences conditional on her new evidence. (*Equal*)

Equal suffices for the present expository purposes, but will receive discussion in Chapter 6, which argues that it is ambiguous in an important way.

Pictorially described, conditionalization proceeds as follows: the agent removes his credence from all of the doxastic possibilities containing not-E, and redistributes that credence over the E-possibilities; this will add credence both to the remaining P and to some not-P possibilities. What is crucial is that the redistribution follow in proportion to the credence in P which the P-possibilities enjoyed before the agent became certain that E. Straight conditionalization proceeds by plugging the agent's values for P and E into the following formula:

[73] As we are about to see, if an agent's credence in E changes instead to 0, he cannot conditionalize on E because the equation by which conditionalization proceeds will be undefined if E takes on a 0 value. But this is not a problem, for the agent can instead run the equation by conditionalizing on not-E, in which, if he is coherent, his credence will be 1.

Bayes's Theorem[74]

$$\mathrm{Cr}(P|E) = \frac{\mathrm{Cr}(E)\mathrm{Cr}(E|P)}{\mathrm{Cr}(E)}, \qquad \text{where } \mathrm{Cr}(E) > 0$$

Bayes's Theorem gives a way to derive the value of a conditional credence in the event that an agent's unconditional credences are defined.

A stronger version of this formula adds to the caveat that Cr(P|E) is undefined when Cr(E) is not greater than 0; that is, it *defines* conditional credences in terms of unconditional credences. But there are significant counterexamples to a definition of conditional credences in terms of unconditional ones.[75] Further, there are a few reasons instead to take conditional credences as basic, not least of which is that unconditional credences are themselves conditional, namely on the tautology. To maintain neutrality about whether conditional credences are defined in terms of unconditional credences or vice versa (though I am personally convinced by Hájek (2003) that conditional credences are basic), I opt for the weaker formulation of Bayes's Theorem which I have given. Note that, unless otherwise specified, by 'credences' I mean both unconditional and conditional credences.

Bayes's Theorem incorporates three factors. The first, 'Cr(P)', is the prior credence which the agent assigns to P – that is, prior to receiving evidence E. The second, Cr(E|P), is her prior credence in E when P is certain, also called the *likelihood* of the hypothesis on the evidence. The third, 1/Cr(E), is the inverse of the agent's credence in her evidence proposition. Since Bayes's Theorem is used only when the agent's credence in her evidence is certain, this value will always be 1. Mathematically, this has the effect of *renormalizing* the agent's credence distribution; that is, of ensuring that the agent's credences in P and not-P always sum to 1 – which, recall, is the standard maximum of the normed measure given by the Normality axiom.

---

[74] Here is a proof. Take the definition of conditional probability,
(1) Cr (P|E) = Cr (P & E) / Cr (E).
Multiply both sides by Cr (E) to get
(2) Cr (P|E) * Cr (E) = Cr (P & E).
Again, by the definition of conditional probability,
(3) Cr (E|P) = Cr (P & E) / Cr (P).
Multiply both sides by Cr (P) to get
(4) Cr (E|P) * Cr (P) = Cr (P & E)
From (2.) and (4.) we know that both 'Cr (P | E) * Cr (E)' and 'Cr (E | P) * Cr (P)' are equal to 'Cr (P & E)' and therefore
(5) Cr (P|E) * Cr (E) = Cr (E|P) * Cr (P)
Divide both sides by Cr (B) to get
Cr (P|E) = Cr (E|P) * Cr (P) / Cr (E)
and this is Bayes's theorem.

[75] Hájek (2003) provides a litany of counterexamples, some of which I discuss in Chapter 5.

As for the caveat ('where Cr(E) > 0'), it makes explicit that, because it is impossible to divide by zero, there is no defined value for Cr(P|E) when Cr(E) is not greater than zero. Hence for purely mathematical reasons, conditionalization is possible only when the caveat is met. This caveat will play an important role in the development of my theory, due to a common but false assumption – to be discussed shortly – about how to interpret it.

Bayes's Theorem gives only the rule for updating in the special case where the agent's credence in E changes to 1 on receipt of *e* as evidence. But an agent's evidence may take on *any* new value except for 0. A proposition E in which the agent's credence is less than 1 (but greater than 0) is typically called *uncertain evidence*.[76] How should updates on uncertain evidence be modelled? Straight conditionalization will not do the trick. The reason is, the agent's prior credence in her evidence (that is, the value by which the product of the likelihood and the prior is divided) is not 1, and so the agent's credences will not renormalize after updating if she tries to perform straight conditionalization.

For this reason there is Jeffrey conditionalization (named for its inventor, Richard Jeffrey)[77]. Jeffrey conditionalization proceeds by applying a formula called 'Probability Kinematics', which is a generalization of Bayes's Theorem, for updating on uncertain evidence:

Probability Kinematics[78]

$$New\ Cr(H) = \sum_{i=1}^{n} old\ Cr(H|E_i) new(E_i)$$

where each E_i is a subjective probability value for a proposition E, such that the total E_is sum to 1. Thus if the agent acquires, as evidence, a new credence in E of *x*, where *x* is less than 1, she inputs *x* for E_i and 1-*x* for E_{ii}.

Implementing the Probability Kinematics formula amounts to conditionalizing twice – once on the agent's new credence in E and once on her corresponding new credence in not-E – and weighting the outcomes by the value of the agent's credences in E and in not-E, respectively. In other words, first she must plug a value of 1 for E into Bayes's Theorem; then she must conditionalize in accord with her conditional credence for (P|E), weighting the outcome by the value of her credence in E; then she must plug a value of 1 for not-E into Bayes's Theorem, conditionalize in accord with her conditional credence for (P|not-E), weighting the outcome by the value of her credence in not-E;

[76] This name may be misleading, suggesting to the impressionable that evidence itself is uncertain and is thus a mental attitude. Far from it. What is uncertain is the agent's credence *in* her evidence, the latter being a proposition.
[77] See Jeffrey (1992, 2004).
[78] Jeffrey (1992), p. 54.

finally, she must add the weighted outcomes, to arrive at her posterior credence in P on receipt of uncertain evidence E.[79]

Why think that Jeffrey conditionalization is needed in addition to straight conditionalization? One reason may derive from the logical positivist claim that evidence must be subject to expression in an 'observation language' common to all inquirers,[80] and from the observation that languages are too poor to express the range of observations that might be made. Enter Jeffrey conditionalization on uncertain evidence: Where the only relevant options for evidence are (for example) *The liquid turned blue*, *The liquid turned green*, and *The liquid turned purple*, none of which can be given certain credence because in fact the liquid turned something like blue-ish-purple-green, Jeffrey conditionalization allows agents to update on a logical partition, in which each of these propositions receives an uncertain credence all of which sum to 1.[81] However, Chapter 3 argues against the view, which I call 'Data Bayesianism'; I argue that evidence is not limited to what can be expressed in a common 'observation language'; an agent's evidence, rather, consists in any proposition from which one may rationally engage in theoretical reasoning.[82]

A better reason to employ Jeffrey Conditionalization – indeed, the my reason – is that it is needed to accommodate rational updates on evidence in which the agent has a credence of less than 1. Some think that there is no such evidence but Chapter 3, Part II will argue (among other things) that there is. Although this does not entail that Jeffrey conditionalization is the rational way to respond to such evidence, there is to date no developed alternative which employs Bayes's Theorem as effectively. Rational updating is characterized by conditionalization, either straight or Jeffrey.

Conclusion

There are volumes of arguments for Bayesian Statics, many of which also support Weak Bayesian Statics, and for Bayesian Updating. These include synchronic and diachronic 'Dutch-book arguments', arguments from decision theory to the effect that agents who are expected-utility maximizers obey these rules, and self-styled 'epistemic' or 'non-pragmatic' defenses employing the notion of truth utility.[83] There are also numerous criticisms. Some claim that Bayesian Statics and

[79] This process proceeds automatically, as described by the mathematics.
[80] For discussion of this view see Hoyningen-Huene (1993), 213-214.
[81] Indeed, Jeffrey's proposal is developed for credences defined over sentences, not propositions (see his 2004).
[82] The Bayesian notion of evidence is a term of art, as I have said, and so I am not making any claims about any natural-language notions of evidence.
[83] For Dutch-book arguments, see Lewis (1999) and Teller (1973). For decision-theoretic arguments see Jeffrey (1983) and Maher (1993). For 'non-pragmatic' arguments, to the effect that conditionalization maximizes expected truth (or epistemic) utility, see Joyce (1998, 2009) and Greaves and Wallace (2006). For an argument

even Weak Bayesian Statics over-idealize.[84] Some claim that Bayesian Updating does not give a necessary condition for diachronic rationality.[85] I won't discuss these arguments here. My purpose in the following is, employing Weak Bayesian Statics, to defend Bayesian Updating against two specific forms of criticism, given in Chapters 6 and 7, which result from implicit yet false assumptions about how it is to be implemented.[86]

I have presented the main features of the Bayesian epistemological framework, and some of the ways in which Expansive Bayesianism develops it. But there is both room and need for much elaboration. As the Introduction emphasized, Bayesianism itself is the foundation and skeleton of a house, but to be a theory of rationality it must be built upon. Much building has been done already in the literature, though (I'll argue) in a haphazard fashion much of which needs to be undone so that something systematic and conducive to a good theory of rationality can replace it. Since its inception Bayesian epistemology has been deployed in a broad spectrum of applications, including and especially scientific reasoning and the philosophy of science, and more recently in epistemology. The promise shown by Bayesianism in this plethora of applications is part of what has driven Bayesian epistemology forward. But in being applied in this or that context it has inevitably acquired a residue of implicit assumptions. These assumptions must be made explicit and examined on their own merit, and retained only if they further the creation of a true theory of rationality. The following chapters begin this work of de- and re-construction.

---

that agents with 'integrity' conditionalize when they have committed to doing so, see van Fraassen (1984, 1995). See Hájek (2008) for an insightful discussion of the limitations of these arguments in justifying Bayesian Updating. For arguments for Jeffrey Conditionalization, see van Fraassen (1989, Chapter 13), van Fraassen (1986), Armendt (1980), and Skyrms (1987).

[84] See e.g. Foley (1993) and Plantinga (1993).

[85] See Williamson (2000), Levi (1967), and 'Sleeping Beauty'-type arguments discussed (e.g.) in Stalnaker (2008), Lewis (1983), Elga (2000).

[86] Jeffrey Conditionalization is subject to a completely different form of criticism which stems from the fact that it is *non-commutative*, i.e., that the order in which an agent receives and updates on a sequence of uncertain evidential propositions will affect the posterior credences she assigns on the basis of that evidence. I won't discuss this issue here. See Domotor (1980) and Döring (1999) for arguments that updating on uncertain evidence should be commutative and hence that Jeffrey Conditionalization is defective, but also see Lange (2000) for an argument that such updating need not be commutative, and, moreover, that at least sometimes it should not be. On the basis of Lange's argument I do not think that the commutativity issue is a problem for Jeffrey Conditionalization.

3. Building on the Foundation: Evidential Input

Introduction

Bayesianism gives the scaffolding for a theory of rationality. That scaffolding leaves much to be filled in, especially about the nature of the most basic elements which it treats. Some crucial ones are credence and the input step (in which an agent assigns a credence to her evidence). To set the stage for my own Bayesian theory, this chapter argues in favor of some of their key characteristics. The claims established here give a picture of rationality as a property of sentient agents capable of entertaining and engaging with the objects of their credences, as opposed to a mechanical picture of rational agents as robots programmed to crunch numeric data.

Part I follows through on a consequence of a claim introduced in Chapter 1, that the ideally rational agent need not grasp all propositions. The consequence, which was introduced in Chapter 2 but will receive more thorough discussion here, is that one of the following claims must be denied: Ideally rational agents have credences defined over all propositions (*Complete*), or having a credence in a proposition entails grasping that proposition (*Grasping*). I argue against Complete and in favor of Grasping.

Part II establishes some Expansive Bayesian claims about evidential input. It argues that there are rational constraints on two out of three aspects of the input step. These aspects are the evidential proposition to which the agent assigns an input credence on the basis of an experience, and the value of that credence. The aspect not subject to rational constraint is the kind of representational

experience on which an input credence may legitimately be based. Expansive Bayesianism thus runs directly counter to a common Bayesian view, which holds that the representational experience is constrained, the evidence proposition is constrained in an unduly limiting way, and that the input credence is not constrained at all. I argue that all of these claims are incorrect, and that Expansive Bayesianism sets experience free to influence the agent's reasoning as it should.

Part III argues, first, that the coming to grasp of a new proposition provides a source of evidence which has the potential to radically alter an agent's representation of reality, and second, that coming to grasp new propositions is a much more regular occurrence than one might think. These insights pave the way for the following chapters to begin making my case that Bayesianism is a flexible and powerful theory of updating when divested of certain burdensome assumptions with which it is often, unnecessarily, associated.

Part I: Credence and Propositional Grasping

This section establishes two claims which are not typically articulated in Bayesian literature but whose falsehood is more in keeping with the ethos of that literature than their truth is. I shall argue however that the two claims in question are true. Establishing them will create room for the insights of Expansive Bayesianism, which run counter to the established ethos in refreshing ways.

To begin, recall that Chapter 1 argues that this claim is false:

Being ideally rational entails grasping all propositions. (*Rational Grasping*)

The idea is that propositional grasping is an epistemic good which is distinct from rationality: an agent can be ideally rational and yet there may be propositions which she does not grasp.

The negation of Rational Grasping has far-reaching consequences. For one, it opens the possibility that an ideally rational agent can come to grasp *new* propositions, which can then, as Part III of this chapter argues, feature as completely novel evidence for an agent, which she could not have anticipated receiving before coming to grasp the proposition. Another consequence of Not-Rational Grasping is that at least one of the following claims is false:

Being ideally rational entails having credences defined over all propositions. (*Complete*)

or

Having a credence in a proposition *p* entails grasping *p*. (*Grasping*)

Complete pertains to rationality itself, whereas Grasping pertains to credences.

That Not-Rational Grasping entails the negation of at least one of these claims can most easily be seen by noting that Rational Grasping and Grasping entail Complete: That ideal rationality requires grasping all propositions, and that having a credence in a proposition requires grasping it, entail that ideal rationality also requires grasping all propositions. Hence, to turn around the entailment, Not-Rational Grasping entails either Not-Complete or Not-Grasping. Since I endorse Not-Rational Grasping, I must choose whether to deny Complete or Grasping.

We'll look at both claims and at two competing pictures of rational credence change which motivate each. I'll argue that Complete is false and that Grasping is true. The choice of Grasping over Complete will set the tone for the rest of the theory of rationality to be developed in these pages. There is more at stake than the issue of whether to reject Complete or Grasping. What is at stake is the underlying picture of rationality. What would an ideally rational agent be like of whom Complete is true and Grasping false? Such an agent would be no more than a machine programmed to crunch numeric input into numeric output, regardless of whether she (or, rather, *it*) has any understanding of the things which those numbers express. But I submit that rationality, even ideal rationality, is not mechanical. It is a property which can only be had by sentient beings who have the capacity to engage mentally with reality by way of how they represent it to themselves.

Complete is False

What case might explicitly be made for endorsing Complete? Here is an argument:

1. Being ideally rational entails having credences defined over all logical truths and all contradictions. (*Complete Logical*)

2. Having a credence in a proposition *p* entails grasping *p*. (*Grasping*)

3. If an ideal agent's credences are defined over a proposition *p*, then her credences are also defined over each proposition which is a constituent *p*. (*Proper Constituent*)

4. Therefore, being ideally rational entails having credences defined over all propositions. (*Complete*)

What are we to make of this argument? Premise 2, Grasping, is true, and I'll motivate it shortly. Premise 3 also has at least a prima facie plausibility. I think it is false, since I think that a rational

agent can be certain that it will either rain or not rain, without having any idea how to apportion the probabilities more finely.[87] But let us spot premise 3 to the proponent of Complete.

That leaves us with premise 1, Complete Logical. This claim seems unshakeable from a Bayesian point of view. After all, don't the probability axioms entail it? For, as they are often stated, they range over all propositions:

Statement of Probability Axioms, with Domain of Propositions Left Out

For all propositions, Cr(p) ≥ 0. (*Non-negativity*)

If p is a tautology, then Cr(p) = 1. (*Normality*)

For any countable sequence of propositions p_1, …, p_n, …, such that any two propositions in the sequence are mutually exclusive,

$$Cr\left(\bigvee_i p_i\right) = \sum_i Cr(p_i)$$

(*Countable Additivity*)

Those who state the axioms this way do us a disservice, as Chapter 2 argued. The axioms should, and do as I have stated them,[88] hold *only* for the propositions over which an agent's credences are in fact defined. Even in the case of physical and statistical probabilities, this quantifier should be explicitly restricted, since these applications of the probability calculus typically do not deal with unlimited domains either. As Chapter 1 argues, there is no correlation between rationality and the number of credences one has. Rationality pertains merely to how one manages those credences. Hence the axiom should be stated as pertaining only to the particular set of propositions in question, as Chapter 2 has done:

Probability Axioms With Domain of Propositions Explicit

For all propositions p which are elements of C, Cr(p) ≥ 0. (*Non-negativity*)

If p is a tautology and is an element of C, then Cr(p) = 1. (*Normality*)

[87] There might be something desirable from the point of view of having certainty in truths to have a credence rather than not, but even this is not clear, since a high credence in a falsehood is arguably worse, from this perspective, than no credence in a falsehood.

[88] See also Eagle (2011) and Howson and Urbach (1993).

For any countable sequence of propositions p_1, …, p_n, …, such that any two propositions in the sequence are mutually exclusive, and each p_i is an element of C,

$$Cr\left(\bigvee_i p_i\right) = \sum_i Cr(p_i)$$

(*Countable Additivity*)

Restricting the quantifier in the probability axioms to the propositions in the intended domain deflates one reason for endorsing Complete Logical: ideally rational agents *need not* have credences defined over all propositions, at least not for any reasons having to do with the probability axioms. There are surely other considerations to be advanced in favor of Complete Logical, but I leave them to its supporters to develop, and take it to be false.

Are there any other reasons to endorse Complete? I can't think of any, and I can think of plenty of reasons to deny it. One is that Complete entails that it is always more rational to have some credence in a proposition than to have none. But this claim is surely falsified by a situation in which an agent does not grasp the proposition in question. Rationality cannot require, sometimes or ever, that an agent have a subjective probability value for a proposition she cannot even entertain. Yet if Complete is true (and Rational Grasping is false, as I am assuming), then rationality will demand sometimes having credences in ungrasped propositions, since Complete and Not-Rational Grasping entail that Grasping – premise 2 – is false.

But Grasping is true, and will become important later in these pages. Let's spend a moment motivating it. Recall from Chapter 1 that an agent grasps a proposition *p* only if *p* represents for her a way things could be, independently of any credences she may have in *p*. Recall also that and having an experience as of *p* entails grasping *p*.[89]

Grasping seems uncontroversial.[90] But some might deny it. Call credences in propositions which one does not grasp *obtuse credences*. Grasping denies that there are obtuse credences. One might defend obtuse credences on these grounds: An agent with credences of *x* in *p* and of *y* in *q* is, as long as she maintains these credences, committed to a credence of *xy* in (*p* ^ *q*) (assuming that *p* and *q* are independent), given the probability axioms. This is so regardless of whether she grasps (*p* ^ *q*). (She may not grasp it if it is, say, too long or complex for her to entertain, even if she does grasp *p* and *q* individually.) It is unclear what the attitude of commitment might be, but let us charitably assume that it is something like a disposition, in the appropriate circumstances, either to have the credences which follow by the probability axioms from the credences one already has or to reject the latter, where appropriate circumstances are those in which one deduces those consequences or

[89] This is not to say that grasping is necessarily conscious.
[90] Williamson (2000), Chapter 9, argues that grasping *p* is necessary for having *p* as evidence.

is told them by a reliable source who has done so. An obtuse credence is what one has until one comes to grasp the proposition in question, in which case the credence becomes non-obtuse. In this case, every rational agent would, in virtue of grasping any proposition at all, have an obtuse credence at least in every logical truth, and probably also in many contingent propositions which he does not grasp. This is an interesting line of thought, but it does not go far enough: a commitment to have a credence is not a credence. Additional motivation would be needed to show that a commitment deserves to be called a credence, too.

The proponent of obtuse credences will suggest the following apparent counterexample, in which the agent really seems to have a credence in an ungrasped proposition: There is a complex astrophysical hypothesis H which Biff does not grasp when it is presented to him,[91] but Biff's mother tells him that H is *x* probable. So Biff goes to the local casino, where there is a roulette wheel with ten areas on which the spinner might rest. One area is named 'H', and each other area is given the name of some other proposition (A, B, and so forth) which is a logical alternative to H. The wheel is set to land on the true proposition. Biff is not aware of which propositions any of these names stands for, including H (possibly he does not grasp any of them), but he is certain, on his mother's epistemic authority, that the spinner will with probability *x* land on H. He thus bets in accord with his mother's subjective probability value, because he is certain that she is an epistemic authority on such matters (she told him that too). Biff can bet on H without grasping H. The best way to explain Biff's behavior, says the proponent of obtuse credences, is that Biff has an obtuse credence of *x* in H.

But I submit that there is a better way to explain his behavior: by positing a credence not in H but in the higher-order proposition *'H' is true* (which, on my fine-grained understanding of propositions, is a different proposition than H, even though they are equivalent). The casino scenario is thus no counterexample to Grasping, which I maintain is true.

Rational Grasping, which says that being ideally rational entails grasping all propositions, is false, and hence either Complete is true or Grasping is true, but not both. I've argued that Complete is false: ideally rational agents need not have credences defined over all propositions; and that Grasping is true: having a credence in a proposition entails grasping that proposition. This sets the stage for a theory according to which rationality is a property possessed not by number crunching machines but by sentient agents who understand the objects of their credence. How an agent represents reality to herself is important for rationality. The following section develops this picture by defending a somewhat unorthodox view of what it is for a rational agent to receive evidential input.

[91] That is, when it is presented to him in a language in which he is competent.

Part II: Evidence and the Input Step

This section establishes some Expansive Bayesian claims about the input step, which is the step of the update in which the agent assigns a credence to her evidence. The input step contains three possible objects of rational constraint: it might constrain (a) the kind of representational experience which can form the basis for the assignment of an input credence to the agent's evidence; (b) the particular proposition to which an agent assigns an input credence: why one proposition and not another? (c) the value of the input credence itself. A common Bayesian view, which I'll call *Data Bayesianism*, says (a) that the kind of representational experience is subject to constraint, and (b) that so is the proposition itself, but (c) that the input value is not. I'll argue that Data Bayesianism is wrong on all three counts: my Expansive Bayesian view holds (a) that there are no constraints on the permitted type of representational experience, (b) that Data Bayesian's constraint on the evidential proposition is wrong but that there may be others, and (c) that there are rational constraints on the value assigned to the evidence proposition. First let's dissect the input step itself.

Updating in Two Steps

Recall from the previous chapter that updating proceeds in two logically (if not temporally) distinct steps: (1) the input step, in which the agent assigns a credence to her evidence proposition, and (2) the adjustment step, in which she changes any of her other credences which need to be changed to ensure that they are (weakly) coherent with the credence which she has newly assigned to her evidence. Bayesian Updating, the claim that the only rational way to update is by conditionalizing, applies to the adjustment step only. It says that, once the agent has assigned a credence to her evidence, this step must proceed by setting her post-update credences equal to the value of her pre-update credences, conditional on her evidence (or by performing the analogous Jeffrey-conditionalization move in the case of uncertain evidence). What makes a theory of rationality distinctly Bayesian, then, is just what it says about the rational response to evidence *once the agent has already assigned an input value to it*. Such a theory is free to say anything about what makes for rational input in the first place. As part of my Expansive Bayesian theory of rationality, I'll stake some claims about the input step which will advance my account later on.

The input step is so called because in it a newly changed credal value is 'input' into the agent's doxastic system. Bayesians call the input step the assigning of credence to *evidence*. Recall from Chapter 1 that 'evidence' is here a term of art intended to designate the proposition to which the agent assigns an input credence so as to usher in an adjustment. It is not intended to map onto the ordinary language term 'evidence', nor onto any notion of evidence employed in coarse-grained epistemology. Indeed, we should not expect the notion of evidence in fine-grained epistemology to be transferrable either from coarse-grained epistemology or from ordinary language, since fine-grained epistemological categories are *sui generis*. 'Evidence' is, however, an apt theoretical term, since Bayesian evidence does retain a few features of many coarse-grained accounts, such as propositionality, and since it performs a function analogous to the function which the notion of evidence plays in everyday English, at least when ordinary English is used to talk about rationally legitimating or mandating a change of opinion.

Are there normative constraints on the input process? Milne (2003) articulates a view that is not unusual among Bayesians in saying that "the details of the story of how [an agent comes to assign an input credence to an evidential proposition] are of no moment" (p. 283), by which I take him to mean that rationality says nothing one way or the other. The arguments to follow tell strongly against this view.[92]

In fact the input step has room for three different normative constraints: (a) one on which kind of representational experience can form the basis for the assignment of credence to evidence; (b) another on the particular evidential propositions which receive new credences on the basis of the experience; and (c) another on the value which the agent assigns to the evidential proposition. I'll argue that (a) other than being a representational experience, there is no constraint on which type of experience is an appropriate candidate for precipitating an input step. This view may sound extreme, but coupled with the second view I'll defend, that (b) the evidential proposition which arises from the experience *is* subject to rational constraint (though not the one endorsed by many Data Bayesians), and that so is (c) the value which the agent assigns to her evidence proposition, my view is not extreme at all. Any type of representational experience can prompt the assigning of a credence to evidence, but there are cases in which not just any input credence will do.

[92] Williamson (2000), Chapter 9, and Dunn (2010), Chapter 2, offer some reasons for why Bayesian epistemology needs a more robust account of evidential input.

(1) Can only some types of experience provide evidence but not others?

Under what circumstances may an agent rationally assign an input credence to an evidence proposition? Bayesianism itself says nothing on this front and is compatible with views on which agents may input credences for no reason at all, because of a knock on the head, or because they feel like it or have an itch. I take it however that a rational agent assigns an input credence because her view of what reality is or could be like has changed. After all, reasoning with evidence is reasoning about what the world is like. Hence I suggest a minimal necessary condition on rational input:

> Any input credence must be assigned on the basis of a *representational experience*. (*The Representational Condition*)

A *representational experience* is an experience which represents things as being thus-and-so: it depicts, describes, or is more generally "about", what is or could be the case.[93] A representational experience brings it about that one or more propositions become salient to the agent, but these propositions might be other than those which the experience represents as being the case.[94] (More on this in the next section.)

Are there any other conditions, or is the representational content condition also sufficient for an experience to provide evidence? I'll argue that it is sufficient. But why would anyone think that it is not? Those I called 'Data Bayesians' endorse the claim that

> An agent can rationally assign input credences only on the basis of a representational experience which is also sensory. (*Sensory Experience*)[95]

A sensory experience is an experience mediated by any of the agent's five senses or by proprioception (if this has a distinctive source). The idea is that evidence can only be provided by an agent's 'data', for emphasis 'hard data', deriving from experiences of color changes in litmus paper, numbers on screens, blood on carpets, and so forth. The motivation behind Sensory Experience

[93] My account is neutral on the issue of whether experiences are intrinsically representational, in the sense of having veridicality conditions, though I am sympathetic to the claim that they do.

[94] That is, the evidential propositions made salient by an experience might not differ from any veridicality conditions which the experience may have.

[95] Data Bayesians stand in the tradition of Quine (1969), according to whom an agent's evidence is no more and no less than the 'stimulation of his sensory receptors' (p. 269), and Russell (1951[1917]), according to whom, "[i]f such objects [as molecules, atoms, and electrons] are to be verified, it must be solely through their relation to sense-data" (108). Jeffrey (2004) is a present-day Data Bayesian.

might be the thought that sensory evidence is the only kind which is in some sense intersubjectively available. Even if this motivation is relevant in certain experimental scenarios (which is itself dubious), it is no good for a more general theory of updating of the sort pursued here. Moreover, there is at least one important counterexample to Sensory Experience.

The counterexample is given by a particular vehicle of representational content which legitimately affects agents' theoretical reasoning yet cannot be accounted for by Sensory Experience. It is what I'll call a *mental experience*. I have in mind the sorts of events that occur in an agent's mind which are not directly imposed by sensory stimuli. They occur when she introspects, imagines, turns puzzles or equations over and over in her mind, grapples intellectually, or attempts mentally to appropriate new theoretical distinctions. Mental experiences are traditionally described by appeal to the suggestive notion of intellectual 'perception'.[96] Descartes (1641) took 'clear and distinct perception' to be veridical, but mental experiences need not be. They are analogous to sensory seemings in that the way they represent things as being might not be the way things are. That an agent has an experience as of A's entailing B does not itself entail that A entails B. An epistemology which allows mental experiences to be evidence stands in direct contrast to the mechanical Data View, since it is premised upon a picture of agents as having internal intellectual lives rather than merely as being subject to the "stimulation of their sensory receptors" (Quine, 1969, p. 269).

Mental experiences can legitimate the assigning of new input credences, even though they do not work through the agent's sensory apparatus. The proponent of Sensory Experience would be hard pressed to deny that an intellectual seeming of a mathematical proposition, which (if veridical) enables an agent to deduce a conclusion which she could not previously deduce, provides evidence for the agent for that conclusion. Non-deductive reasoning works similarly: an agent can ponder a quantum physical equation without grasping it, and, on coming to grasp it and undergoing an associated intellectual seeming as of the proposition expressed by the equation, recognize that it strongly confirms a theory which he had previously dismissed as untenable.[97] Surely the intellectual seeming provides the basis for evidential input.

I thus reject Sensory Experience. Instead, I endorse:

> An agent must assign input credences on the basis of a representational experience, but nothing further constrains the type of experience which this must be. (*Representational Experience*)

[96] See e.g. Descartes (1641), Wolterstorff (1996), and Plantinga (2000).
[97] Chapter 6 gives a Bayesian account of this phenomenon.

Representational Experience says that *any* experience which is a vehicle of representational content can legitimately influence an agent's epistemic reasoning.[98]

One might object that Representational Experience is too liberal. Doesn't it threaten to undermine the project of rationally constraining updates on evidence? Bayesian Updating painstakingly mandates that the adjustment step proceed only by conditionalization, but of what value is this if the agent may assign input credences on the basis of any sort of representational experience at all? What, for example, about a representational experience brought on by a vivid daydream or a hallucinogen? Surely an agent would not be rational in assigning input credences on the basis of such experiences.

My response is this: I don't think that there are any epistemically privileged ways of being presented with representational content. Epistemic anarchy can be prevented by placing rational constraints on either or both of the other components of the input step: the evidential proposition which is assigned a new credence on the basis of the experience, or the value of the credence assigned. A vivid daydream might be representational and hence, by Representational Experience, is a legitimate candidate for giving rise to an evidential proposition. But which proposition that may be is itself limited, as is the value which the agent may assign to it. An agent may daydream that she is onstage winning an Oscar for a screenplay she fantasizes about writing. But this daydream arguably cannot rationally deliver up the proposition *My arms are made of spaghetti*, which (given her background evidence) is not remotely evidentially related to Oscar-winning. A permitted proposition might be *The stage lights which I'm imagining are bright*. This proposition arguably cannot rationally receive a low input credence, but instead must, if the imagined lights are bright, receive a rather high input credence. With restrictions such as these there is no reason to cry 'epistemic anarchy'.

The next section discusses constraints on the input proposition, rejecting one offered by the Data Bayesian but suggesting that there might be others, and the next section argues for a constraint on the value which the agent assigns to the input proposition. Although any representational experience can legitimate the assigning of a credence to evidence, this is far from the end of the story about rational constraints on input.

[98] There are likely many other varieties of representational experience than sensory and mental; some might not even be perceptual in either the sensory or 'mental' senses. Candidates include phenomena such as blindsight, subconscious monitoring, and so forth; I do not want to exclude any of these forms of representational experience, but will not discuss them here.

(2) Rational Constraints on the Evidence Proposition

This section discusses and rejects a restriction which Data Bayesians impose on the evidence proposition, and tentatively advances another. The Data Bayesian restriction is this:

> An agent's evidence consists only in propositions about how things seem to that agent. (*Seemings*)[99]

Propositions about how things seem, such as *It seems that there is a red butterfly on your nose*, are to be contrasted with propositions about how things are independently of one's experiences, such as *There is a red butterfly on your nose*. On this view, if I have an experience as of a red butterfly landing on your nose, then the only rational way for this experience to form the basis of a high credence that there is one is to make an inductive step from *It seems that there is a red butterfly on your nose*.[100]

Why endorse Seemings? I'll examine and debunk two motivations. One might be given by a view about how experiences work: that they only give rise to propositions about how things seem. Thus if the propositions delivered up by an experience provide an agent with her evidence, the agent will only have propositions about seemings to work with anyway. Here is an argument along these lines:

Argument from Seemings Descriptive

1. For any rational agent, her representational experiences will only present her with propositions about how things seem, as opposed to propositions about how things are independently of her experiences. (*Seemings Descriptive*)
2. An agent must assign input credences on the basis of a representational experience, but nothing further constrains the type of experience which this must be. (Representational Experience)
3. Therefore, an agent's evidence consists only in propositions about how things seem. (Seemings)

What are we to make of this argument? Premise 2 was defended in the previous section. But Premise 1, Seemings Descriptive, is false.

One might endorse Seemings Descriptive if one is attracted to a sense-data theory of experience, according to which agents do not (directly) experience physical objects, but rather mental entities

[99] The denial of Seemings plays an important role in Chapter 7, which gives an Expansive Bayesian account of rational conversion.

[100] Seemings is probably espoused by Brewer (1999), and a view of evidence which endorses Seemings and Premise 3 of the argument to follow (the claim that evidence is given by agents' representational experiences) is espoused by C. I. Lewis (1947) and, under some interpretations, David Lewis (1996).

called 'sense-data'.[101] If one already thinks that experiences are mediated by sense-data, one might be kindly disposed toward the intuition that the propositions rationally delivered up by experiences are *about* seemings (in the form of sense-data), rather than about external reality. However, any affinity between Seemings Descriptive and sense-data theory is merely superficial. First, sense-data theory does not entail that the only propositions which arise from representational experiences are propositions about seemings. For even if, as sense-data theory says, seemings are the only objects of *awareness*, propositions about seemings need not be the only rational objects of the doxastic attitudes which agents form on the basis of what they are aware of. Second, it is hard to see how sense-data theory would even make it plausible to suppose that representational experiences only give rise to propositions about seemings. Sense-data theory does not provide a motivation to endorse Seemings Descriptive. But this is just as well for the proponent of Seemings Descriptive, since association with sense-data theory is not to my mind a particular theoretical virtue.

More importantly, Seemings Descriptive gets the psychology of credence-formation badly wrong. Typically, an experience as of seeing the kettle on the counter when I turn on the kitchen light brings me to have a high credence that the kettle is on the counter, not that it seems as if the kettle is on the counter. The Data Bayesian might concede that my credence ultimately rests at *The kettle is on the counter*, but he will insist that it winds up there by way of an inductive step from *It seems as if the kettle is on the counter*. Admittedly, this could be so, but why think it is? A much simpler account of the situation is that things are as I experience them: I experience forming a high credence in a proposition about the kettle, and not about any kettle seemings. The Data Bayesian owes us an error theory explaining why agents typically do not experience their credence formation as involving propositions about seemings.[102] I conclude that Seemings Descriptive is false, and hence the Argument from Seemings Descriptive is unsound.

But if considerations about how experiences really work do not support Seemings, the Data Bayesian can simply respond: So much worse for the way experiences really work! What is important are their *normative* features, and only propositions about seemings are appropriate to provide evidence from experiences. This thought might derive from a classical epistemological

[101] See e.g. Russell (1951 [1917]).

[102] This is not to say that representational experiences *never* give rise exclusively to seemings propositions. If it was antecedently very implausible that something be the case (say, that the cat starts speaking Russian), then an experience as if that thing is the case might not give rise to a proposition about reality at all, but only to one about seemings (as in: *It seems that the cat is speaking Russian*). But surely even experiences as of antecedently implausible occurrences can give rise to propositions about more than just the way things seem. An experience as of a winged monkey may well present me with the proposition *There is a winged monkey*, even if it also presents me with the proposition *It seems as if there is a winged monkey*. But forming credences in seemings-propositions and not in propositions about external reality is surely the exception rather than the rule, and surely does not occur in mundane circumstances such as when I have an experience as of the kettle's being on the counter.

foundationalism, according to which the rock-bottom level of justification for one's doxastic attitudes – be it other doxastic attitudes, evidence, or something else – must enjoy some special epistemic status, such as being true or infallible. When the foundation has some such status, conditions are optimal for the doxastic attitudes based on it to enjoy a similarly high status too. Seemings might be advanced as a condition on evidence which confers epistemic status upon the credences which one adopts as the result of reasoning from that evidence. Seemings would ostensibly ensure that the propositions which form the agent's epistemic foundation – her evidence – are those about which she cannot be mistaken. Plausibly, the only candidates for this role, other than tautologies, are propositions pertaining to the agent's experiences. If so, then Seemings would appear to isolate a plausible epistemic foundation for 'data' for agents to reason from.

Here is an argument along these lines for Seemings:

Argument from Evidential Certainty

1. For any proposition *p* presented to a rational agent by a representational experience, she assigns credence 1 to *p*. (*Evidential Certainty*)
2. Rational agents invest credence 1 in propositions about how things seem to them. (*Luminosity*)
3. An agent must assign input credences on the basis of a representational experience, but nothing further constrains the type of experience which this must be. (Representational Experience)
4. Therefore, only propositions about how thing seem may be used as evidence. (Seemings)

Premise 3 is true; the previous section endorsed it. Premise 2, 'Luminosity', is false. First, Williamson (2000) argues convincingly that agents are not always certain about the content of their experiences. One might respond that the case with ideally rational agents may be different; perhaps Williamson's considerations do not apply there. But there is little reason to think that ideal rationality is different in this regard. Considering how strong an intellectual commitment certainty is, it surely can't be considered less than ideally rational for an agent to be reticent about bestowing it, even about her own experiences. Thus there may be scenarios in which a rational agent is certain of no proposition of the form *I am having such-and-such an experience*. She may have a credence of .5 or of .9999, but in neither case will Data Bayesianism permit her to update on this evidence. Thus in a case in which the agent forms an uncertain credence in a proposition *p* as the result of a representational experience, but does not form any certain credence at all, Data Bayesianism cannot accommodate an update on *p*. Premise 2 is false.

The falsehood of Premise 2 is enough to undermine this argument for Seemings. But even if it were true, the argument would fail on account of Premise 1, Evidential Certainty. This premise says that, if an agent has an experience as of a shining green disc coursing through the sky, she must

become certain of any proposition which that experience makes salient to her, such as *It seems as if a shining green disc is coursing through the sky*, or *A shining green disc is coursing through the sky*. Evidential Certainty yields the result that rational agents only ever have credence 1 in their evidence[103]. Of course, the identification of credence 1 with the attitude of certainty can be questioned, but I am assuming for present purposes that this identity holds.

Evidential Certainty is a throwback to an era in which straight conditionalization – that is, conditionalization on a proposition to which one assigns a credence of 1 – was the only game in town for Bayesian updating. The Bayesian model only functioned with a strict application of Bayes's Theorem; Jeffrey conditionalization, the point of which was to accommodate uncertain evidence, had not yet been invented. Some Bayesians thought that, because their model was correct, evidence could only rationally be allotted a credence of 1. Other more modest Bayesians thought that their model was limited: that there may be rational assignments of uncertain credence to evidence, but only updates by agents with a credence of 1 in their evidence could be modelled. Bayesians of both stripes were moved to endorse Evidential Certainty. And since it would be problematic to become certain of *any* proposition arising from an experience (such as *There is a butterfly on your nose*), the Bayesian was inclined to claim that only privileged propositions – those about seemings – could be evidence. But Evidential Certainty is outdated. Jeffrey conditionalization provides a perfectly Bayesian way to accommodate uncertain evidence, as Chapter 2 has shown; it is Bayesian because it amounts to conditionalizing twice on each member of a logical partition and merely weighting the results. Thus those Bayesians who concluded, on the basis of their limited model, that evidence *could* not be certain, are challenged to revise their view on the basis of the enhanced (Jeffrey) model (and perhaps, for next time, not to base their views about rationality on limited mathematical models to begin with); and those who held out for allowing that there might be uncertain evidence which their model was too simple to accommodate were vindicated. Evidential Certainty is false.

Although the arguments I have considered do not exhaust the ways in which one might defend Seemings, I take Seemings to have been put to rest for now. The agent's evidence proposition is not restricted to propositions about how things seem to her.

However, I have not excluded the possibility that there are other constraints on the evidence proposition. Indeed, it makes sense that there would be. I will not advance any such constraints here. I'll suggest merely that some propositions do seem to be more aptly associated with a given experience than other propositions are. For example, it seems apt, on having an experience as of rain in the garden, to assign input credences to such propositions as *It is raining in the garden*; *the sky is gray*; *the leaves are wet*, and so forth, and not to such propositions as *the moon is made of*

[103] Whether Evidential Certainty is true has important consequences for the way in which the adjustment step proceeds, as we'll see in the next section.

blue cheese or *Peanut butter and jelly taste good together*. Recall also our Oscar and spaghetti example from above. Particular criteria for the aptness of an evidential proposition given some experience will be difficult to spell out precisely, and I won't try to here.[104] Suffice it to say that the apparent aptness or in-aptness of some propositions given some experiences suggests that there is a rational distinction in play.[105]

Restricting which evidential propositions may rationally be generated by a representational experience gives us a response to the objection posed to Representational Experience, above. Recall that Representational Experience was objected to on the grounds that it is allegedly too lenient, allowing for epistemic anarchy in the input step even as it mandates conditionalization in the adjustment step. But constraining the proposition which serves as evidence permits us to respond by saying that, although any representational experience may provide evidence, there are sometimes very tight restrictions on which propositions it provides evidence for. Without saying more about the kinds of restrictions there may be, I'll leave this thought about restricting evidence propositions here, and turn my focus to a third form of constraint on the input step: namely, on the value of the credence which the agent assigns to her new evidence.

(3) Rational Constraints on the Input Credence

This section argues that rationality constrains the input value which the agent assigns to her evidence proposition. Bayesian Updating already entails some such restrictions. In particular, when an agent has acquired new evidence *e* as the result of a previous conditionalizing update, her credence in *e* must take on the value of her credence in *e* conditional on whatever evidence she last received. Bayesian Updating not only constrains, but mandates, an agent's input credence for *e*.

But what about when the agent's input is a credence in a proposition *e* which is not the result of a previous conditionalizing update? In this case there is no formula for determining any particular value which her credence in *e* may or must rationally have. Are input credences acquired in this way subject to rational constraint? Some Bayesians respond with an unequivocal 'no'. This is Jeffrey's own view, in the face of critics who say that Jeffrey conditionalization is deficient for not providing a 'rule' determining the particular value which an agent's credence in her new evidence[106] takes on as

[104] There may be different sorts of constraints on which propositions can be evidence as the result of an experience. For example, there may be 'local' constraints, on the relation between a single experience and any proposition(s) to which it legitimately gives rise; or there may be 'global' constraints, on the relation between the total experiences an agent has had and any proposition(s) which their sum makes salient to her.

[105] Brewer (1999) provides a holistic account of rationality (or 'reason') as incorporating the propositions made salient by experience, which may lend itself nicely to the present account.

[106] In particular (and in Jeffrey's parlance), the values taken on by each cell in the partition over the possible evidence propositions.

the result of an experience.[107] Such criticism focuses on the acquiring of *uncertain* evidence from experience,[108] but there is no reason to stop there. The same problems arise for justifying the assigning of *certain* credence to new evidence: why should an experience generate credences of any particular value, be it 1 or less than 1, at all? Jeffrey's response (and that of many) has been to deny the need for such a rational justification. He claims that

> The value which an agent assigns to her evidence proposition as the result of a representational experience is not subject to rational evaluation. (*A-Rational Input*)

A-Rational Input is false.

Note first that A-Rational Input is committed to two normative asymmetries. One is an asymmetry between input derived by conditionalization, which is subject to rational evaluation, and input derived by experience, which is not. This asymmetry is not so bad, since the proponent of A-Rational Input may hold that input derived from conditionalization is rationally evaluable *qua* posterior credence, but not *qua* input value. *Qua* input value, it, too, is a-rational.

But there is a more important normative asymmetry to which the proponent of A-Rational Input is committed; namely, one between the two steps in the updating process. Whereas it takes the adjustment step to be rationally evaluable (saying that it must proceed by conditionalization), it denies that the input step is. Why subscribe to this asymmetry? The only relevant difference between these steps, as far as I can tell, is that the adjustment step is rationally governed by a formula, Bayes's Theorem, whereas the input step is governed by no formula. But this difference may be reason in itself for those who endorse A-Rational Input. That is, they might presuppose that

> The value of a credence can be subject to rational evaluation only if it can be derived from input by the application of a logical or numeric rule. (*Justification by Rule*)

Endorsing Justification by Rule is understandable, given the success of conditionalization and given that credences are, after all, numeric. If this assumption underlies the dialectic between Jeffrey and his critics, then Jeffrey's response seems the only way out: since no such rule can be given for deriving the value of credences generated by experience, then this value must be a-rational.

---

[107] Gabbay, Woods, and Hartmann (2011), for example, write that Jeffrey conditionalization "is incomplete in a very important way. Without some supplementary rule telling us which partition an experience bears on, and what the probabilities on that partition the observation warrants, we cannot apply the rule. Jeffrey Conditionalization needs a partition and a distribution over it as inputs, and we haven't been hold how to select these inputs" (p. 502).

[108] For discussion and an attempted re-configuration of Jeffrey conditionalization in light of this issue (and the non-commutativity issue; see Chapter 2), see Field (1978). But see Garber (1980), who argues that Field's solution has unacceptable consequences.

But we do not need to endorse Justification by Rule. I don't endorse it, and hence may also deny A-Rational Input. Denying A-Rational Input gives us a second response to the objection posed to Representational Experience, above. The objection, recall, was that Representational Experience is allegedly too lenient, allowing for epistemic anarchy in the input step while mandating conditionalization in the adjustment step. One response was to restrict the proposition which serves as evidence. The second response, denying A-Rational Input, lets us respond that, although any representational experience may provide evidence, there are restrictions on what values the agent may assign to that evidence. I submit that some values, *when coupled with certain experiences*, are rational whereas others are not. For example, it is surely irrational, in response to an experience as of rain and with no other relevant background evidence, to form a very low credence that it is raining. Sometimes, even, the uniquely rational credence to assign to an evidence proposition may be *no different* from the credence which one had in that proposition before it was newly presented to one. This might be the case for example when an agent is newly presented with a proposition by free association or by being compelled by brainwashers to stare for days at sentences expressing it. Although any proposition may be evidence, agents may not respond to their evidence in any way they please.[109]

The considerations advanced in Part II of this chapter have important consequences: There are *two* rationally evaluable steps in the updating process. In order of conceptual (and perhaps also temporal) priority, these are assigning an input credence, and adjusting for it. Until now, Bayesianism has focused on the properties of and norms governing the latter.[110] I suggest that this has perpetuated a strong tendency for Bayesian theories to locate all of their normative eggs in the basket of conditionalization. But diversification is needed: I think that equal focus must be devoted to the acquiring of evidence. That input values cannot be made precise by way of a numeric or logical rule does not matter. The input step makes up a full half of the updating process. If this is so, then the acquiring of evidence is fertile ground for the positing of epistemic norms, and those who favor robust normativity have much to cultivate in exploring this process.

[109] As in the case of constraints on propositions arising from experience (see previous footnote), in the case of constraints on the value of an input credence, different constraints may arise 'locally', on a single experience / credence pair, or 'globally', on a credence given the totality of the agent's experiences up to that point as well as her background evidence. I won't spell out this suggestion but note it to give a taste of the rich variety of input constraints there may be.

[110] To substantiate this claim I call attention to the vast literature on conditionalization cited in the rest of this dissertation, in comparison with the paucity of resources which come up in searches for 'Bayesian input', 'evidence', and so forth.

Part III: Newly Grasped Propositions as Evidence

The final part of this chapter builds on the claims established thus far, bringing further detail to the picture I am establishing of rationality as a feature of sentient agents with mental lives. From Representational Experience, the claim that an agent may only assign input credences on the basis of a representational experience but that nothing further constrains the type of experience this must be, it follows that:

A newly grasped proposition can be evidence for an agent. (*Newly Grasped*)

Newly Grasped follows from Representational Experience because coming to grasp a proposition for the first time is a representational experience, in which the newly grasped proposition is presented to the agent. Newly Grasped does not say that *any* newly grasped proposition may be evidence. I suggested above that there may be rational restrictions on which proposition(s) an experience makes salient to an agent. Newly Grasped says merely that, within these limits, a proposition which an agent comes newly to grasp is a viable candidate for being evidence for him. Since coming to grasp a new proposition is an experience which has representational content (that is, the content of the newly grasped proposition itself), doing so is a representational experience and hence an evidential experience.

How may an agent come to grasp a totally new proposition? One way is for his conceptual capacity to expand: he may learn to draw distinctions which enable him to think at a finer grain.[111] An agent from the tropics, who is only familiar with one notion of snow, may learn to distinguish twenty different kinds of snow by having them pointed out to him: 'One kind of snow is like that', 'Another kind of snow is like that', and so forth.[112]

Let's call new evidence which the agent did not grasp before her evidential experience *undefined evidence*, shorthand for 'evidence which the agent did not grasp (and hence over which her credences were not defined) before she acquired it'. Since agents can come to grasp new propositions, it makes sense to suppose that they can also undergo what Chapter 1 called 'C-set expansions': the set of propositions over which their credences are defined can expand or, more precisely, become a bigger set.[113]

[111] See van Fraassen (2002) for discussion of this phenomenon.
[112] Being linguistically competent is not a necessary condition for entertaining propositions; take the dog which grasps that its mistress is leaving it food. See Williamson (2000) for discussion.
[113] Of course an agent may come to grasp a new proposition without any C-set expansion occurring: she may suspend judgment from the proposition.

From the points made in the first two sections of this chapter, Newly Grasped might look obvious. But it is an important claim to put onto the Bayesian table explicitly. The reason is that it can, and will in subsequent chapters, be brought to bear in showing Bayesianism to be a much more flexible theory of rational updating than many give it credit for.[114]

But more basically and prosaically, Newly Grasped is an important component of an account of how agents acquire evidence by having representational experiences. The coming to grasp of new propositions, we'll see shortly, is an important and regular feature of agents' epistemic lives. Showing how this is so will provide us with a more sophisticated understanding of the receipt of evidence from sensory as well as mental sources, in addition to putting an important tool in place to show, in subsequent chapters, just how flexible a Bayesian theory of updating can be.

The rest of this chapter will prepare the way for the discussions in the following chapters, by arguing that

Coming to grasp new propositions is an ordinary feature of agents' epistemic lives. (*Common*)

Common is not a claim about the statistical probability which some suitably average agent has of coming to grasp some number of new propositions over her lifetime. I do not intend it to be anything like that precise. I offer it merely to emphasize this point: Updates on evidence which was defined before the evidential experience are a canonical form of rational update, by which I mean that they should be modelled by any ideally rational agent. But updates on defined evidence are not the only canonical form of update: *updates on undefined evidence are canonical too*. The reason is that a wide variety of evidential experiences in fact include the coming to grasp of a new proposition.

Why bother arguing for Common? Because one might think that it is only in exceptional cases that agents come to grasp a new proposition, as when an agent comes to entertain a new theory which she had never before considered. On the whole, one might think the grasping of new propositions is the exception rather than the rule: they are not a very important part of agents' everyday credence formation. But this blasé view is erroneous. We need a more sophisticated understanding of how evidence arises from sensory and mental experience.

[114] Complaints about the inflexibility of Bayesianism will be made precise later. But here is a brief preview of two claims to which Bayesianism is often accused of being committed. One is: 'Once an agent's credence in a proposition *p* has reached 1 (and thus her credence in not-*p* has reached 0), the agent's credences in *p* and in not-*p* cannot rationally change' (*Unrevisable*). The other is: 'Conversions are irrational'. Some readers will think these claims are true, but Chapters 5 and 6 argue, respectively, that they are false. I mention them here only to advertise that the claim 'Newly Grasped' will factor into my argument against them.

The rest of this chapter argues that undefined-evidence experiences are not limited to exceptional and high-level academic theorizing. They also occur when agents have complex yet run-of-the mill sensory experiences. These two kinds of evidential experience in which the agent comes to grasp a new proposition will be brought to bear later in showing off the power of a Bayesian theory of updating.

Mental Experiences, Including Grasping New Theories

The coming to grasp of new propositions has already received discussion among Bayesians, in the special case in which the newly grasped proposition is a theory which the agent had not understood before the evidential experience in which she comes to grasp it (more below). But the grasping of a new theory is a special case of what we called above a 'mental experience'. Such experiences do not arise from sensory stimuli but rather from the agent's turning over thoughts in his mind. The theoretical 'turning over in one's mind' can be aided by the direction of a teacher, as Socrates aided Meno.[115] The 'aha!' moment is one in which the agent experiences an intellectual phenomenology. We saw above that mental experiences can deliver up evidence for deductive conclusions, as in the case of Meno, or for inductive conclusions, as when an agent grasps a piece of data, say a complex equation, which confirms a theory.

One might think that mental experiences are relatively infrequent, occurring only at the highest abstractions of theoretical research. But this view is mistaken. Meno grasped a simple new theoretical proposition, and presumably so does every agent undergoing theoretical education at any level, from primary-schoolers to professors. Any time an agent learns to draw a new conceptual distinction, her doxastic space is carved up more finely than it was before, adding a new proposition (and its negation) to the set over which her credences are already defined. Or any time she learns a new concept which is not referred to even ambiguously by any she is already capable of deploying, completely new doxastic possibilities open up where there were none before. It makes sense to think that the grasping of new propositions is a regular feature in the doxastic lives of most agents. Grasping new propositions continues as long as one's education does, which is to say, potentially for a lifetime.

The grasping of new theories has received discussion among Bayesians because, in some cases, it is regarded as a problem for Bayesianism. The problem is said to arise when the agent has a combined credence of 1 in what she takes to be mutually exclusive and exhaustive alternative theories; call them $\{H_1, \ldots, H_n\}$. Say that she comes to grasp and assign a credence to yet another

[115] See Plato (1961).

theory, H_{n+1}, which is a rival to $\{H_1, \ldots, H_n\}$. She may do this by reading a scientific journal which introduces and propounds the novel and revolutionary H_{n+1}. If she is to assign a credence to H_{n+1}, this credence must come from somewhere. Since her credences already form a partition over $\{H_1, \ldots, H_n\}$, in the sense that they sum to 1, there seems no way to assign a credence to H_{n+1}: she can only do so by way of a non-conditionalizing update. But such an update would of course be irrational. How the agent can rationally assign a credence to H_{n+1} is known as the 'problem of new theories'. [116]

The Bayesian response to the problem of new theories is to suggest that the rational agent holds open the possibility that there are logical possibilities on which she has not missed out; she does this by maintaining a non-zero credence in a proposition, called a 'reserve hypothesis', such as *There is a theory which I have not yet grasped*. The idea is that, on having an evidential experience which delivers up H_{n+1} as evidence, she can simply assign to it some of the credence she had allotted to *There is a theory which I have not yet grasped*. The problem of new theories is thus limited in scope, since it only arises when the agent lacks a reserve hypothesis. There is little reason to think that agents often lack reserve hypotheses – only the most dogmatic agent takes there to be no propositions she has not yet entertained. And there is less reason to think that it is rational to lack one. But Chapter 5 will show that Expansive Bayesianism can accommodate even the claim that agents who unfortunately do lack reserve hypotheses can nonetheless rationally form credences in newly grasped propositions.

I conclude that mental experiences, including but not limited to the grasping of new theories, are rather common evidential experiences.

Complicated Sensory Experience

Having a mental experience is one way to come to grasp a new proposition. But even more often, arguably, agents come to grasp new propositions by having run-of-the-mill sensory experiences.

To begin, note that our imaginative capacities are limited. Many sensory experiences have phenomenological characters which exceed what an agent is capable of fully imagining in advance of the experience, where an agent *fully imagines* an experience just in case she duplicate its precise character in her mind. We may also say that an agent fully imagines an experience just in case she grasps a proposition which has as its content every single aspect of that experience. Such a proposition need not be expressible by the agent in her language; indeed many such propositions are probably not.[117] If the agent does not grasp such a complex and specific proposition in advance

[116] For discussion see Maher (1995), Earman (1992), and Shimony (1970).

[117] I remain neutral about whether complicated sensory experiences can have non-conceptual content, but for various essays on this issue see Gunther (2003); see also Peackocke (1992).

of the experience which delivers it up as evidence, then she cannot have a credence in it before the experience occurs, since (as I have argued) agents must grasp propositions in order to have credences in them.

Thus whenever an agent has a sensory experience which she could not fully imagine before having it, she comes to grasp a new proposition. The proposition is something along the lines of *Things are like that*.[118] Take the proposition *Niagara Falls is like that*, which the Belgian monk Louis Hennepin, who has visited the Falls, grasps but which King Philip of Spain, who has not, does not grasp.[119] There are at least three possible reasons why the King does not grasp *Niagara Falls is like that*: (1) It refers to an aspect of reality which falls completely outside of his space of doxastic possibilities, in the sense that it does not entail any propositions which he already grasps. (2) It is too specific for him to grasp, even if he does grasp a more general proposition which is entailed by it. (3) It has too much content for him to grasp before he is brought to do so by an experience. Some detail on each reason follows.

(1) King Philip might not grasp *Niagara Falls is like that* because it contains elements which fall outside of his imaginative capacity altogether – he might not even have more general categories under which they are subsumed. If for example instead of being the King of Spain he were the king of an arid planet such as Mars, an experience of Niagara Falls would acquaint him for the first time with the notion of liquid, expanding his doxastic possibility space. Arguably young children come to grasp completely new propositions on a regular basis in this way, since they become more epistemically sophisticated day by day or even, in the very early stages of life, hour by hour. Experiences of type (1) are likely an everyday occurrence for a young child, and decrease in frequency as he grows older. But arguably they do not cease altogether, especially for agents who travel often and expose themselves to new ways of thinking.

(2) The second reason why King Philip – from Spain this time, not from Mars – might fail to grasp *Niagara Falls is like that* is that this proposition is too fine-grained: it contains elements which he cannot distinguish from other propositions which might describe Niagara Falls in counterfactual but similar possible scenarios. For example, before he experiences Niagara Falls, he grasps *Purple is like that*, but after his experience of the manifold refracting colors of the water droplets, which introduce distinctions in purple shades which he had not previously encountered, he grasps *Purple is like that*, Purple is like that**, Purple is like that****, and so forth. Since *Purple is like that* is entailed

[118] Helpful discussion of such cases can be found in Brewer (1999), who suggests that they be expressed by the sentence 'This is thus'; McDowell (1994); and Williamson (2000). For a critique of this approach see Christensen (1992).

[119] For simplicity let's assume that, were King Philip to visit the Falls, he would be able to grasp *Niagara Falls is like that* even when 'that' refers to Father Louis's Falls phenomenology, since we shall assume that it does not include any phenomenology of uniquely experiencing the Falls *from Louis's perspective*.

by the latter three more specific propositions, the experience brings the King to grasp propositions which are entailed by a proposition which he already grasps – namely, *Purple is like that* – but which, until the experience, were too specific for him to distinguish. I am thus relying on a principle to the effect that grasping a proposition P which is entailed by Q does not entail grasping Q, since an agent's conceptual apparatus might be sophisticated enough to grasp P but not Q. I take the Niagara Falls example to indicate that this principle is plausible.

Experiences of type (2) are arguably slightly more frequent than experiences of type (1), since it is arguably more common to learn to parse concepts which one already has than to come across completely new ones; and, similarly, likely occur frequently among young children and less frequently among more mature epistemic agents, though can be brought about by travel and exposure to new ideas and sensations.

(3) The third reason why the King might fail, before seeing the Falls, to grasp *Niagara Falls is like that* is that it has too *much* content. Even if the King could grasp each of a sequence of propositions describing every single aspect of the experience, he might not have the epistemic capacity to grasp their conjunction all at once without the help of the experience which imposes it on his awareness. Among many propositions which an experience of Niagara Falls might present to the King, let us zero in on two. One is

E1 *There is a large and misty waterfall before me.*

Before the experience, the King already had a credence defined over E1, and, since there is little counterevidence, his experience induces him to raise his credence in E1 significantly.

However, E1 is a very general proposition, in that it comprises a rather large set of mutually exclusive logical possibilities (such as *E1 and it is sunny, E1 and it is not sunny,* and so forth). There are infinitely many more specific propositions which entail E1 yet which are too finely parsed for the King, a limited agent, to have individual credences in each of them. The various aspects of the experience are too numerous to be contained in any proposition which he can fully imagine and hence grasp before the experience occurs. E1 only contains the most general aspects of the experience. But the Falls experience delivers up a second proposition, *Niagara Falls is like that*, which he did not grasp before his experience, where 'that' refers to the full Falls phenomenology.[120] Even though the King grasped E1 before his experience, he did not grasp *Niagara Falls is like that* – and this is so even though the latter entails the more general E1. Here I am relying on a second principle, that grasping two propositions *P*, *Q* does not entail grasping *P and Q*. It would stand to

[120] Let's ignore that the referent of 'that' may change as the agent's memory fades.

reason that such a principle might be operative, since human cognizers – even potentially ideally rational ones – have limited mental processing capacities.

Whereas experiences of types (1) and (2) arguably occur with decreasing frequency as one grows up (but never stop entirely), experiences of type (3), I submit, are everyday occurrences no matter how old an agent is. The reason is that our senses take in more information than we can hold in our imaginations without having the experiences in question. Thus these three ways of coming to grasp new propositions are not limited to highly unusual experiences. Arguably they arise as a result of all or most sensory experiences. This phenomenon is not typically discussed in Bayesian literature, but arguably it should be, since C-set expansions brought about by complex sensory experience are, if the Falls example is apt, a regular feature of agents' epistemic lives.

Future chapters will show that recognizing this helps Bayesianism respond to some hard objections against it.

Conclusion

Part III has argued that the coming to grasp of new evidential propositions is a regular rather than exceptional feature of agents' doxastic lives, and that it occurs by way of sensory as well as mental experiences. That being ideally rational does not require grasping all propositions entails that coming to grasp new ones is a feature of ideal rationality as much as of non-ideal rationality. However, it is not just the coming to grasp of new propositions, but the assigning of credence to them, which makes them important for rational updating, since they must have a credence assigned to them in order for the input step to proceed. A theory of rational updating must therefore account not only for the coming to grasp but also for the assigning of credences to newly grasped propositions, and the adjusting of one's other credences in light of the input value.

Expansive Bayesianism accounts for the rationality of both of these things, and in doing so shows Bayesianism to be a perhaps surprisingly flexible theory. Bayesianism has been the target of objections, which I'll call *liberationist*, to the effect that it is too strict to allow for some important rational phenomena such as radical conversions, the possibility of updating to one of two different posterior credences, and more broadly the exercise of certain forms of epistemic virtue which are had by epistemic agents and not by robots. These criticisms will be countered in the chapters to follow. The present chapter has established the basis for responding to them, by giving an initial picture of rationality as pertaining essentially to the engagement of sentient beings with the representational content which they can entertain. The theory of updating to be built up from here will silence liberationist critiques of Bayesianism.

Before addressing the criticisms, the following chapter outlines and criticizes a prominent liberationist view intended to replace Bayesianism, the 'voluntarist' epistemology of Bas van Fraassen.

4. Van Fraassen's Theory of Updating: A Critique

Introduction

Bayesian epistemology is subject to a genre of criticism we may call 'liberationism'. The concerns voiced from this corner are many, but they fall under a general heading which accuses Bayesian epistemology of being too *strict*. In the tradition of Kuhn (1962), Feyerabend (1975), and James (1956), they regard rationality as involving epistemic virtues such as artistry, creativity, spontaneity, good doxastic intentions, the freedom of the agent to determine her own doxastic life, and a certain lightness of touch concerning the importance of evidence.

Liberationists come in many stripes, but they have in common their rejection of Evidentialism (see Chapter 1), the claim that

> The receipt of new evidence (i) rationally mandates an update, and (ii) is the only thing which rationally permits one. (*Evidentialism*)

The clause which liberationists deny is (ii); they think that non-evidential considerations may rationally factor into the way in which an agent updates her credences, at least in some circumstances. Bas van Fraassen is a liberationist. He criticizes the dictum of the famous evidentialist, Clifford (1879), that "it is wrong, always, everywhere, and for everyone to believe anything on insufficient evidence" (1989, 172). He endorses the liberationism of the pragmatist James (1948), "which recognized rationally allowed leaps, far beyond the solid security of previous opinion-plus-evidence so far" (1989, 173). Van Fraassen asks, "Could we rationally have, in any other

way, come to our present high opinion of the theories of Darwin, Einstein, and Bohr, whose empirical implications stretch throughout all dimensions of world history, past and future?" (*ibid*.).[121]

One of van Fraassen's reasons for denying clause (ii) of Evidentialism is that he subscribes to a strong form of doxastic voluntarism; he thinks that agents have voluntary control over the way in which they update their credences, and that rationality permits an agent, in some circumstances, to update by choice instead of by following her new evidence, where an agent follows her new evidence by conditionalizing. He speaks of allowing for "true epistemic *decisions*" (1984, 170, italics in original), asserts that "[b]elief is a matter of the will" (1984, 164), and posits, as we'll see, a robust notion of doxastic commitment which is expressed in the agent's choice about how to update. I will examine van Fraassen's doxastic voluntarism more closely at the end of this chapter; it will prove to be a major problem with his theory of updating.

Two of van Fraassen's main allegations are that Bayesianism cannot account for rational conversion from one theory to another, and that it is incompatible with permissiveness about updating (the claim that there is more than one rational posterior one might adopt). His arguments for these claims will be countered, respectively, in Chapters 6 and 7. The present chapter criticizes the positive theory of rational updating with which van Fraassen aims to replace Bayesianism. His theory of updating is part of an epistemological package, called 'voluntarism', which includes but is not limited to the diachronic aspects of his epistemology which will be discussed here. For simplicity of expression, I use the term 'voluntarism' synonymously with 'van Fraassen's theory of rational updating', and hence in a narrower way than van Fraassen himself does.

This chapter exegetes voluntarism, in the sense just given.[122] Aside from our concern with defending Bayesian Updating, exegeting the diachronic elements of van Fraassen's epistemology is worthwhile in itself. There is a rich trove of ideas about updating which have only seen partial discussion in the literature.[123] Van Fraassen is well known for two claims which on the face of it may seem to compete: one is that it is irrational to update by following a rule of reasoning other than conditionalization (such as inference to the best explanation), and the other is that it can be rational to update by following no rules at all, not even conditionalization. His argument for the first claim has been subject to much discussion, but comparatively little has been said about the second. These claims do hold together logically, but this chapter argues that they do not gel into a viable theory of

[121] See also van Fraassen (1984, 163).

[122] I thus do not discuss the synchronic elements of voluntarism, including the Reflection principle (for which see van Fraassen, 1984 and 1995).

[123] Psillos (2007) offers a clear and helpful exegesis of voluntarism as a whole, but since van Fraassen himself does not separate out his claims about diachronic rationality in a particularly systematic fashion, neither does Psillos. Separating out van Fraassen's diachronic claims and commitments is one goal of the present exegesis.

updating. Fortunately, however, van Fraassen's criticisms of Bayesianism (we'll see in the final two chapters) do not stand, and hence voluntarism is not needed as an alternative to Bayesianism.

Van Fraassen Against Bayesian Updating

Van Fraassen develops his voluntarist theory of rational updating within the framework of an otherwise Bayesian-looking epistemology: he maintains not just Bayesian Statics, the claim that coherence is necessary for synchronic rationality, but the stronger claim, which we may call *Orthodox Statics*, that coherence suffices for it too. And he holds that conditionalization has a crucial normative role to play in rational updating. He differs from Bayesianism only in denying Bayesian Updating, the claim that

> Conditionalization is the only rational way for an agent to update her credences. (*Bayesian Updating*)

Van Fraassen claims that there are instead other ways to update rationally than by conditionalizing. His theory of rational updating can thus be criticized by those who dislike its Bayesian aspects or by those who dislike its non-Bayesian aspects. Being a Bayesian, I shall critique it for the latter.

Although van Fraassen thinks that there is something right about Bayesian Updating, he denies the claim as it stands. Bayesian Updating is equivalent to the conjunction of two claims, of which he rejects the first and endorses the second:

> The only rational way for an agent to update her credences is in accord with a rationally permitted rule. (*Rulebound*)
>
> The only rationally permitted rule by which agents may update is conditionalization. (*Unique Rule*)[124]

A *rule* is a function from input of a particular form to a single, determinate output. Conditionalization is the rule which one follows by redistributing one's credences, on receipt of evidence, in accord with Bayes's Theorem (see Chapter 2); it takes as input an agent's credence in a proposition newly acquired as evidence, together with all of her other credences, and gives posterior credences as

[124] Note that van Fraassen thinks that evidential input can take other forms, such as conditional probabilities, for which other rules – such as 'Infomin' (1989, chapter 13) – are permissible; but I think he is wrong in thinking that there is any form of non-propositional input. At any rate, this discussion pertains only to propositional input.

output. To update by a rule is simply, when one receives the evidence, to apply the rule.[125] Bayesian Updating entails Rulebound, since conditionalization is the only permitted form of update and it is a rule; and Bayesian Updating entails Unique Rule, because it denies that agents may update in any other way than by conditionalizing. Together, Rulebound and Unique Rule entail Bayesian Updating.

Van Fraassen denies Rulebound and endorses Unique Rule (1989, Chapter 7). That is, he holds that rationality permits agents to update by doing something other than abiding by an updating rule, and a fortiori other than conditionalizing, but maintains that, if an agent does update by a rule, that rule must be conditionalization.

Since I aim to criticize only the non-Bayesian aspects of van Fraassen's theory of updating, I won't discuss his argument for Unique Rule. As for his two main arguments against Rulebound, they will be countered in Chapters 6 and 7. For a flavor of them, however, we may note that van Fraassen excoriates the Cliffordian evidentialist for being "just like Carnap's robot: his senses bring him propositions that he takes as evidence, and his total response to this consists in *conditionalizing* his present state of opinion on those propositions" (1984, 154), and scoffs at the "methodological cookbook" he thinks Bayesianism prescribes (2002, 89). The thought is that Rulebound prevents the "rational free enterprise of the spirit" (1989, 172). Saving for Chapters 6 and 7 the arguments behind these suggestive remarks, this chapter restricts itself to laying out the full theory of updating with which van Fraassen supplements his denial of Rulebound and to critiquing it on its own merit.

Two Additional Voluntarist Constraints on Updating

Van Fraassen's denial of Unique Rule and his endorsement of Rulebound yield two models of rational updates: (i) conditionalizing updates, and (ii) non-conditionalizing updates.

This picture is sparse. As it stands, nothing regulates the updating of agents who reject conditionalization: they could update in any way that strikes their fancy. Voluntarism has more to say. But I'll argue that it doesn't say enough to make a viable theory of rational updating.

Although voluntarism holds that there are no rules which an agent must follow in order for her update to be rational, it holds that there are some *constraints* by which she must abide. Whereas a rule specifies a single, determinate output value for each input value, a constraint specifies, for each input value, a range of possible output values; another way to put the difference is to say that a rule excludes all possible output values but one, whereas a constraint excludes at least one output value

[125] For present purposes we will ignore questions about cases in which the agent follows the rule by accident, intends to follow the rule but doesn't, and so forth, as well as deeper questions about what it is to follow a rule to begin with. For the present exegetical and critical purposes, these questions are not of paramount importance.

but allows the rest. Van Fraassen posits four constraints on updating, of which we've seen two. One is the denial of Rulebound, *Not-Rulebound*. Not-Rulebound is a constraint because it allows multiple kinds of non-conditionalizing updates, excluding only those which abide by rules other than conditionalization. (Rulebound itself is a rule, since it allows only one kind of update.) Van Fraassen's second constraint on updating is Unique Rule (which he shares with Bayesianism), which is a constraint because it limits the kinds of update one may perform, but not to one single kind.

Van Fraassen gives two more constraints on updating. One is:

> In order permissibly to opt out of conditionalizing one's credence in a proposition *p* on new evidence *e*, one must not, before receiving *e*, be committed to conditionalizing one's credence in *p* on *e*. (*Constraint 3 on Updating: No-Commitment*)

For van Fraassen, *committing* to do *x* is like intending to do *x*, or entering into a contract with oneself to do *x* (1984, 162).[126] What is the object of an agent's commitment? It is: *to conditionalize on receipt of evidence*.[127] A defining feature of van Fraassen's voluntarism, then, is the claim:

> If the rational agent makes a doxastic commitment, that commitment must be to conditionalize on receipt of evidence. (*Commitment*)

The time at which the agent commits to conditionalizing a credence in *p* on evidence *e* (should she choose to commit) is *before* she receives *e* (1989, Chapter 7). If she does commit to conditionalizing, then she is rationally obligated,[128] when she receives *e*, to conditionalize her credence in *p* on it. If, on the other hand, the agent has opted not to commit to conditionalizing her credence in *p* on *e* at the time she receives *e*, she is permitted, on receipt of *e*, to update in any way she chooses – provided that it not be by following some rule other than conditionalization.

Van Fraassen's two models for rational updating break down, then, into submodels:

(ia) conditionalizing updates by agents who are committed to conditionalization at the time of updating;

---

[126] This citation discusses commitment in the context of credence formation and not of deciding whether to commit to an updating rule, but I take it that the same notion of commitment is operative in the case of committing to conditionalization, and van Fraassen doesn't say otherwise.

[127] Van Fraassen also talks about commitment to 'epistemic policies'. Two interpretations are feasible: according to one, 'policy' is a synonym for 'rule'. According to another, a policy is a general description of what the agent intends to do when evidence comes in; on this reading, the having of an updating policy is partly definitive of being an epistemic agent. Since nothing hangs on the notion of policy for my purposes, and in the interest of simplifying matters, I will ignore it and speak only of rules.

[128] That is, she is committed as long as the commitment stands; see below for discussion of whether commitments can be rescinded.

(ib) conditionalizing updates by agents who are not committed to conditionalization at the time of updating;

(ii) non-conditionalizing updates.

The submodels which I'll criticize particularly are (ia) and (ii). Non-conditionalizing updates are not subject to a similar division, for such updates may only rationally be performed by agents who are not committed to conditionalization at the time of updating.[129]

Van Fraassen takes doxastic commitment seriously, pressing home its importance for one's doxastic life by positing a fourth and final constraint on updating:

> An agent may not update in a way which violates any of her doxastic commitments. (*Constraint 4 on Updating: Integrity*)

Epistemic integrity is a virtue which van Fraassen lauds in his (1984, 1989, and 1995). He argues for conditionalization by appeal to a thought experiment about a weather forecaster; he states that he is not offering the more typical Dutch-book argument for conditionalization, but rather an appeal to the "forecaster's integrity, his role as a professional" (1995, p. 16), and he makes it clear that he approves the "commitment to stand behind one's own commitments" (1984, pp. 163-164). A guiding tenet of voluntarism is that epistemic agents are really *agents*, with the freedom but also the responsibility which this brings: the freedom to opt out of updating by rule (and thus to exercise the "rational free enterprise of the spirit", 1989, p. 172) but the responsibility to do what one is committed to doing.

Van Fraassen seems to be thinking of a commitment to conditionalize as tantamount to commitments to abide by what has been called a 'paradigm', which is roughly a large-scale theory of reality which one presupposes in pursuing scientific research.[130] For van Fraassen, agents' initial credence distributions seem to be paradigms, and to conditionalize in accord with one's credence distribution seems to be to change one's credence on receipt of evidence in accord with a paradigm. One may not be certain that one's paradigm is the most accurate or theoretically fruitful representation of reality (however these notions are cashed out), but one takes an educated and thoughtful bet on it and steps out in the hope that it will prove effective in learning about the world and in helping one achieve one's scientific aims (1989, Chapter 7). Van Fraassen's idea seems to be that commitment to a paradigm, and following through on that commitment in one's research, puts

[129] Van Fraassen is well known for positing a constraint on credences at a time, called the Reflection Principle (1984, 1995). This yields a constraint on what an agent may rationally commit to. But Reflection has no consequences for how to update if one is not committed.

[130] See Kuhn (1962) for the notion of paradigm most germane to van Fraassen's epistemology.

one in a position to build a more detailed and uniform representation of reality than if one had opted instead not to subscribe to any paradigm in particular. Research has to start somewhere, and the chronically uncommitted agent denies herself a starting point. When one is committed to a paradigm, conditionalization is a way to ensure that one follows through with one's chosen research program, unless or until one makes a responsibly deliberated decision to abandon it. Moreover, commitment to a paradigm ensures that one continue to interpret one's evidence in light of it, for only in doing this, the thought seems to be, can the agent begin to tell whether the paradigm is an accurate or theoretically fruitful one.[131]

The rest of this chapter argues that voluntarism is untenable. But fortunately the theory which it was intended to replace, Bayesianism, does not fall prey to the liberationist critiques which van Fraassen launches at it. This is good for epistemology, since voluntarism, as we are about to see, is not an option.

Against Voluntarism I: Problems with van Fraassen's Notion of Doxastic Commitment

There is a question which van Fraassen does not address explicitly: whether voluntarism permits agents to *rescind* their doxastic commitments and, if so, under what circumstances they may do so. This is an important question: If commitments may rationally be rescinded, then they may be made with a light hand, for the moment they are rescinded, they cease to apply; if commitments may not be rescinded, then they are surely only for the few stout of heart.

Whichever view van Fraassen takes on the rescinding of commitments to conditionalize, however, there is a problem. To see it, take an agent who is committed to conditionalizing, and imagine that his circumstances are of the sort van Fraassen describes (2002, Chapter 3) as looking ripe for a non-conditionalizing update: as evidence comes in, he finds that he needs recourse to ever more auxiliary hypotheses to account for predictions gone wrong, in a cycle that only worsens the more evidence he receives. An onlooker might begin to suspect that this agent's credence distribution is not connected to reality in a way that promotes truth utility (see Chapter 1). There are

[131] Another reason favoring commitment, on this view of matters, might be that it is efficient, enabling agents to exercise discernment in allocating their cognitive resources. Deciding to break from one's paradigm requires thought, energy, time, and no little soul-searching; and this is even before one makes the arduous decision about what credences to replace one's paradigm with. Facing such a decision afresh every time one receives evidence would make the research process slow and painstaking, to say nothing of potentially traumatic. Since abiding by one's paradigm will anyway be the modus operandi in all but the most unusual circumstances (1989, chapter 7), one may as well spare oneself the trouble except in cases in which one suspects that a particular future evidential experience will lead one to new insights in the event that one courageously embraces it.

two mutually exclusive possible cases: Either (1) this agent is permitted to rescind his commitment to conditionalization, or (2) he is not permitted. Both cases are problematic.

The first horn is this: If the agent is permitted to rescind his commitment to conditionalize, then that commitment is worth little. What is the point of committing to conditionalization if the commitment is not rationally binding after one makes it? One may as well not commit to begin with, since anyway it is always permissible to conditionalize, commitment or no. One might defend van Fraassen (if he adopts this horn) by suggesting that the commitment to conditionalize is like an engagement to marry: you can rescind your commitment at any moment before the wedding, but after the priest pronounces you husband and wife – read: after the evidence comes in – it is too late. In itself this suggestion might be workable (we'll come to it in a moment). But this is not how van Fraassen's view works; there is an important disanalogy. In the engagement and marriage case, there are two *different* commitments, made at different times: a commitment, made at the time of the engagement, to marry, and a commitment, made at the wedding, to love and honor until death us do part. It is morally reprehensible to break either commitment, although the consequences to breaking an engagement may be less dire. In the updating case, on the other hand, there is just one commitment: to conditionalize. Whether it is rescinded before or after the evidence comes in, a commitment is a commitment, and it is of little moment of rescindable.

The second horn is this: If the agent may not rationally rescind his commitment to conditionalize, then he is forbidden from abandoning a credence distribution which turns out to be somehow faulty. In effect a two-tier epistemology emerges: at the top are the elite epistemic entrepreneurs who can do as they please because they have not bound themselves with commitments, and at the bottom are the grunts who in their need for epistemic security have sold themselves into bondage to their present credence distributions. Consideration of the grunts reveals that this horn separates into two sub-horns. Either (a) such an agent is fazed by the epistemic trauma of having to come up with ever more auxiliary hypotheses in spite of his best efforts and wishes that he could renege on his commitment so as to rationally update in some non-conditionalizing way, or (b) he is not fazed at all and creates the auxiliary hypotheses as a matter of course.

Let's consider sub-case (a), in which the agent wants to switch credence distributions but is not rationally permitted. He wants to switch them, it makes sense to suppose, because there is another credence distribution which, if updated in accord with it, would garner greater expected truth utility (see Chapter 1) than his present one. But he is not rationally permitted to do so, because he has committed to the latter. According to voluntarism, then, rationality *forbids* this agent from abandoning a credence distribution which, by his own lights, is deficient, and requires that he continue to update in accord with it. I consider this to be an unacceptable result of case (a).

Rationality may have norms which imperfect agents cannot meet. But surely it is not *built into* a theory of ideal rationality that it sometimes be rational for an agent to update in accord with credence distributions which they deem faulty. If the rational way to update is one which the agent himself takes to be irrational, then something has gone badly awry – this situation cannot ever arise in an ideal epistemic situation, which is after all the sort of situation at issue here.

Let's move on to sub-case (b) of the second horn of the committed-agent dilemma. In this subcase, the epistemically enslaved agent is not fazed by the constant need to supplement his hypotheses with auxiliary hypotheses. It is thus hard to see what good the commitment does to begin with. After all, the purpose of a commitment is surely to bind oneself into doing something which one may *not* always be inclined to do. This is why Ulysses is for van Fraassen a telling example of commitment (1995): he has himself bound to the mast because he is aware that he will not *want* to abide by his commitment to sail past the sirens. More prosaically, we typically commit to being faithful for better or worse, or to stick to our diet even when we crave chocolate; but (psychological, dietary, or moral blocks notwithstanding) there is no need to commit, say, to eating chocolate whenever the mood strikes. Such a commitment is pointless.

To summarize the committed-agent dilemma: First, if agents are permitted to renege on their commitments to conditionalize, then such commitments are worthless. Second, if agents are not permitted to renege on their commitments, then those commitments are either (a) irrational to maintain when the agent regrets his commitment, or (b) pointless when he always wants to do what he has committed to doing. Thus in every case in which the committed agent may find himself, the result is deeply problematic.[132]

Van-Fraassen's notion of doxastic commitment is flawed, and voluntarism as it stands should be rejected. But the voluntarist notion of commitment seems an odd one, anyway. Why should dogmatic commitment to conditionalization and absolute epistemic freedom-in-the-moment be the only options? An alternative is this: An agent who does not want to enslave herself to conditionalizing might instead make a more circumspect commitment, namely, to employ conditionalization as her *default* means of updating. The idea is that she commits to conditionalize in the event that no other permitted means of updating strikes her as preferable either before or after she receives her evidence. Conditionalization, notes van Fraassen (1989, 174), is what agents will almost always choose anyway. This commitment to use conditionalization as a default is much weaker than a commitment which binds one when the evidence comes in. It is like an engagement

[132] There is reason to think that van Fraassen opts for the second horn, permitting agents to rescind their commitments to conditionalize: he wants, after all, to convince committed Bayesians that they are mistaken. It makes sense to suppose that he would endorse their rescinding any unfortunate but universal commitments, made by their previously mistaken selves, to conditionalize.

which is not a commitment to marry but a commitment to marry unless one opts out before the wedding.

Van Fraassen would not like this option, because it is presumably not strong enough to prevent agents from succumbing to doxastic temptations that arise in the face of particular evidential experiences. Recall Ulysses, who binds himself to the mast of his ship to prevent himself from sailing to destruction against the sirens' cliffs. Ulysses has committed to his navigated course and commits to preventing himself from diverting from it. The extreme measures are needed, the thought is, because, at the time of commitment, he assigns the highest expected utility to arriving home, and is aware that, on hearing the sirens' song, his utility values will undergo a temporary change which, if he follows through on them, will result in his never arriving home. Van Fraassen seems to worry that a similar peril faces the agent who does not bind herself to conditionalizing. It is committing to *conditionalizing*, rather than conditionalizing if one feels like it, that will save one from straying from one's credence distribution in an ill-advised moment of doxastic temptation.

Van Fraassen's appeal to the siren myth is suggestive. But it is not clear how it applies in the case of credence formation. For van Fraassen explicitly permits agents to decide in the moment of receiving evidence how to update on it (these agents fall under case (ii) above). He does not admonish them for straying from their credence distributions; he instead warmly ascribes to them the epistemic virtues of 'courage', 'spontaneity', and 'creativity'.[133] So it is not that he thinks that there is siren-like epistemic danger in straying from one's credence distribution. The issue really seems to be, for van Fraassen, that there is something fundamentally rational about *either* committing to conditionalization itself (not to some watered down conditionalization-as-default option), *or* to declining to commit at all. It is not clear what that something is, though, and the committed-agent dilemma provides good reason for abandoning the project of spelling it out at all.

Against Voluntarism II: Too Much Freedom, too Little Responsibility

I have argued that the notion of an agent who is committed to conditionalizing (falling under voluntarism's category (ia) above for rational updating) is more problematic than meets the eye. Let's now turn our attention to the agent who is completely uncommitted (who falls under category (ii) above) and hence who is permitted to update as he pleases on receipt of evidence, provided that

[133] For example: "What of the person who says: 'I can envision all of these possible episodes, one and only one of which will come to pass – I do not know now exactly what opinions and expectations I will form in response, I shall in most respects make up my mind then and there, hopeful and confident that I shall proceed both rationally and creatively' – is he not rational?" (1989, 174). He also praises "theoretical innovation and courageous embrace of new hypotheses" (*ibid.*), and the allowing of "room for independence and enterprise in forming our world picture" (1984, 163).

he obey no rule other than conditionalization. There are large problems in claiming that such an agent is rational. We saw that voluntarism claims to strike the right balance between epistemic freedom and epistemic responsibility, according agents the freedom to opt out of updating by a rule while giving them the option of holding themselves responsible to conditionalization. But in the case of uncommitted agents, I submit that instead voluntarism allows for too much freedom while imposing too little responsibility.

Look closely at the voluntarist claims discussed above, and note that, in spite of their array, van Fraassen gives no account of what form or forms an uncommitted update might in fact take. Once the agent has decided against committing to conditionalizing, he may update in whatever way strikes his fancy, as long as he follows no rule in doing so. This seems crazy: since everything is permitted which is not explicitly forbidden (1989, 171), voluntarism permits any number of outlandish updates on evidence. For example, it permits an agent to choose a credence of zero for propositions in which she has, at the time of choosing, a non-zero credence. Surely a theory on which such updates come out as rational cannot be right. But according to voluntarism, any update which the non-committed agent chooses is just as rational as any other. We may call this the *objection from subjectivism about (uncommitted) updates*.

Van Fraassen addresses a similar objection by appealing to the integrity constraint. He points out first that to form a credence is to "take a stand"; it is "akin to commitment, intention" (1989, 179). Then he asks, "what sort of criticism is it to note that I could, without flaunting the bounds of rationality, have taken a different stand? How much pause should that give me? / The fact is, in taking this stand, I knew *that* already." He argues that he would not *now* be rational if he had any credences which differ from the ones which he now has, even though there is a rational process by which, had he chosen it, he could have arrived at them. He appeals to readers to "imagine a friend of laws of nature [the existence of which van Fraassen vehemently denies] who insists that, by my lights, the reality of laws must be as credence-worthy as any other hypothesis which goes beyond my evidence" (179). Van Fraassen contends that he would be irrational to have a high credence in laws of nature – and hence that the objection from subjectivism fails.

But why would he be irrational in now assigning a high credence to the claim that there are laws of nature? The reason, he says, is that the proponent of such laws cannot provide arguments for them which van Fraassen could rationally accept. The reason she could not do so is that van Fraassen's own credences, which include credences about what sorts of arguments for laws of nature would be persuasive, are the fruit of a stand he has taken. He has committed to them, and, according to the integrity constraint, an agent cannot rationally update in a way which violates any of his doxastic commitments. Thus van Fraassen's response to the objection from subjectivism

seems to be that an update to a high credence in laws of nature would be irrational for *him* because of his commitments, even though they can be rational for the agent in question because of hers. But if 'rational credence formation by my lights' and 'rational credence formation by your lights' does not amount to subjectivism about rationality, it is hard to say what does.

I won't argue against subjectivism about rational updating itself; I'll merely suggest that a more robust conception of rational updating is desirable in the case of uncommitted agents – especially given the relatively strict constraints governing committed ones. Voluntarism allots such agents too much doxastic freedom, with too little responsibility to counterbalance it.

Against Voluntarism III: Updating is Not Strongly Voluntary

I just argued that voluntarism succumbs to a form of subjectivism about rationality, on account of not sufficiently constraining rational updates, and hence allows for too much freedom. This section argues that there is a different kind of freedom which voluntarism allows too much of, or rather, which voluntarism presupposes yet which has nothing to do with rationality. It is the psychological freedom to choose between two or more different ways of updating on one's evidence. Voluntarism is false because it presupposes a strong form of doxastic voluntarism. I'll first discuss the form of doxastic voluntarism which van Fraassen requires, and then argue why it is a component neither of rationality in general nor of the rationality governing human agents in particular.

Voluntariness is a property predicated in the first instance of actions. There are two senses of 'voluntary' which might be ascribed to actions. On one sense, which we'll call *weak voluntariness*, an act is voluntary just in case the agent's will is causally efficacious in bringing it about. This is to say that his will plays a causal role, regardless of whether this role is necessary or sufficient; so an act is weakly voluntary even if the agent could not have done otherwise had he so willed. An agent who chooses to stay in a room which he is unaware is locked does so in a weakly voluntary manner. However, that an act is weakly voluntary does not entail that the agent's will suffices for performing it. On the contrary, contextual factors must be in place as well. If the agent is on an island with no buildings and hence no rooms, she cannot voluntarily stay in a room even if she wills to do so. The formation of a credence is weakly voluntary just in case the agent's will is a non-sufficient and non-necessary contributing cause.[134]

But weak voluntariness is not what van Fraassen has in mind. He endorses a stronger notion, which we'll call *strong voluntariness*, according to which an act is voluntary just in case the agent could have done otherwise than perform it. Thus an act is not strongly voluntary *merely* in virtue of

[134] Weak voluntariness is roughly that of Hume's 'spontaneity' (*Treatise*, 1975). Although my account is compatible with robust notions of causation, I am not committed to any.

issuing from the agent's will. Rather, it is strongly voluntary only if she could have performed a different act instead, had circumstances been identical except for the inclination of her will. Hence the agent's remaining in the locked room is not strongly voluntary, even if she chooses to remain there. She must also *be able* to leave the room. The formation of a credence is strongly voluntary just in case the agent could, merely by willing (and carrying through on her will), have formed some other credence in an otherwise identical situation.[135] Van Fraassen needs strong voluntariness because, on his view, agents have real options about how to update: they are able to choose whether to update by conditionalizing, and if they choose not to, they are able to choose how else to update instead.[136]

Any act (or update) which is strongly voluntary is also weakly voluntary, but an act (or update) which is weakly voluntary need not also be strongly voluntary.

We may distinguish four varieties of doxastic voluntarism:

All updating is weakly voluntary. (*All Weakly*)

Some updating is weakly voluntary. (*Some Weakly*)

All updating is strongly voluntary. (*All Strongly*)

Some updating is strongly voluntary. (*Some Strongly*)

All Strongly entails All Weakly (which in turn entails Some Weakly), but there is no entailment from either All Weakly or Some Weakly to All Strongly. Similarly, Some Strongly entails Some Weakly, but there is likewise no entailment from Some Weakly to Some Strongly. Thus anyone who thinks there is any strongly voluntary updating is logically forced to think that there is also weakly voluntary updating, but not vice versa.

We have said that van Fraassen's voluntarism requires strongly voluntary updating. But it is not entirely clear whether he endorses All Strongly or Some Strongly. Let us charitably ascribe to him the less extreme claim, Some Strongly. All Strongly can be summarily rejected, for it entails that even

[135] For van Fraassen, forming a credence is akin to the formation of an intention; I take it that we may think of it as something like a 'doxastic intention', to represent reality to oneself in accord with the credence, which involves endorsing and acting in accord with this representation. If I have formed a credence, van Fraassen says, "I have made a decision, have formed an intention, am committed to a certain stance or program or course of action. There is no direct obligation to anyone else to fulfill this intention, but I have, as it were, entered into a contract with myself. If I express this intention to an audience, then […] I invite them to rely on my integrity and to feel assured that they now have knowledge of a major consideration in all my subsequent deliberation and courses of action" (1984, 162).

[136] That an action or update is strongly voluntary does not entail that the agent's will itself is not causally determined. My accounts of strong and weak voluntariness are neutral with respect to the question of free will.

updates on simple perceptual evidence are strongly voluntary. It entails, for example, that the update from a low credence in *There is a red carpet before me* to a high one, on receipt of the evidence *I am appeared to as if there is a red carpet and have no evidence indicating that there is no red carpet before me*, is something the agent could have chosen, in otherwise identical circumstances, not to perform. Surely such basic perceptual updates are not strongly voluntary. Hence we will not ascribe All Strongly to van Fraassen, but merely Some Strongly.[137] After presenting van Fraassen's particular brand of strong voluntarism, I'll motivate it and then argue that it is false.

First let's briefly discuss the two claims about weak voluntarism. Some Weakly is surely true. All that is needed for it to be true is that some agent somewhere will to form the credence which he in fact forms. One might think that Some Weakly provides inductive support for All Weakly: that some credences are weakly voluntary raises the probability that all credences are weakly voluntary. Even if it does so, however, All Weakly is false. If all updates were weakly voluntary, then no agent would form credences which he does not will to form. But this is surely false. Wanting to be kept in suspense about the end of a film I'm planning to see, I don't want to become confident about how it ends, but can't avoid doing so when you let the ending slip.[138] Or if I want to be certain that I will heal from my illness because I am convinced that being certain will have palliative effects, I will unwillingly become confident that I will not recover if the doctor provides me with enough evidence to that effect.

Having eliminated All Strongly and All Weakly and endorsed Some Weakly, we are left questioning the status of Some Strongly. Van Fraassen assumes it (as far as I can tell he does not argue for it), and we have seen that it provides a major motivation for him to deny Bayesian Updating. For All Strongly says that the agent has a choice, at least sometimes, about how to change credence in response to evidence, whereas Bayesian Updating says that agents may update rationally in only one way, by conditionalizing. But surely, says van Fraassen, rationality permits a more robust exercise of particular forms of intellectual virtue which stem from the agent's will, such as courage, spontaneity, and innovation (for discussion see Chapter 7).

Rationality may permit the exercise of epistemic virtues in the event of any strongly voluntary updating there may be, but van Fraassen's claim is stronger: that rationality *constitutively includes* the ability to exercise strongly voluntary choice in one's credence formation. The diachronic constraints given earlier in this chapter would not make sense otherwise. Recall that van Fraassen's notion of doxastic commitment is a commitment to updating in one way rather than another. Such a commitment would be empty if the committed agent were not able to make good on it: unless she

[137] There is evidence that van Fraassen wants to distance himself from All Strongly, as well; see the discussion in Psillos (2007), 140-141.
[138] This example and other similar ones come from Kelly (2003).

happens to be very lucky in always involuntarily forming the credences she commits to forming, she will continually violate van Fraassen's Integrity constraint, for example. An ideally rational agent must have strongly voluntary choice, on van Fraassen's view. This is an insuperable problem for voluntarism, as I'll now argue.

Is Strong Voluntarism a Feature of Rationality?

Van Fraassen does not motivate Some Strongly, as far as I can tell; he assumes it as part of the bedrock of his aptly named theory. What arguments might one marshal in its favor? One suggestion is that strongly voluntary control is an epistemic good, just as (weak) coherence and conditionalization are. God is as good a candidate as any for being ideally rational, and he arguably does have voluntary control over his credence formation. Strongly voluntary control, one might think, would be of great benefit to an agent who wants to form and manage her credences in keeping with her evidence. Why not build it into a theory of ideal rationality?[139]

There are a few reasons. First, one might indeed think it easier for an agent with strongly voluntary doxastic control to form credences in keeping with her evidence, since she will be able to stop herself from attending insufficiently to her evidence, from being distracted, and from forming credences out of (say) anger or love which contravene her evidence. But strongly voluntary doxastic control could equally *impede* responding appropriately to one's evidence. Voluntary control is only helpful in performing an action if the agent himself wants to perform it. If an agent wills, against his evidence, to form a high credence that his love is requited, or that his rival is a scoundrel, strongly voluntary choice will enable him to do so. Moreover, even the most epistemically virtuous of agents may simply be an incompetent reasoner: if she lacked voluntary choice she might respond magnificently to her evidence, but having to exercise her choice, she invariably makes the wrong decision, often winding up incoherent. Strongly voluntary choice can abet one in following one's evidence, but it can also impede one. It seems to have little to do with rationality.

Moreover, Some Strongly is false, at least of human agents. A theory of rationality should not require that it be true of ideally rational agents. At first blush this response might seem odd. First, there seem to be situations in which human agents *do* exercise strongly voluntary choice about their credences. I'll return to this in a moment. Second, and more fundamentally, why should an agent's psychological abilities affect the rational norms which govern her in the case of voluntarism, but not in the case of Weak Bayesian Statics and Bayesian Updating? These latter norms, after all, make

[139] See another argument against strong voluntarism by Williams (1973), from the claim that belief is constitutively aimed at truth.

requirements which human agents cannot meet. Once we have abandoned an epistemic 'ought implies can' principle, why be selective about what ideal agents should be able to do?

My response is that there is a difference between the Bayesian constraints and Some Strongly. Weak coherence and conditionalization are features of an agent's doxastic attitudes and of how she adjusts those attitudes in refining her representation of reality. Strong voluntarism, by contrast, is a feature of the agent's psychology which has little if anything, conceptually, to do with her doxastic attitudes or with representing reality. If an agent is incoherent (or weakly incoherent), her representation of reality cannot possibly be accurate (however we cash out 'accurate'), but no epistemic consequence follows if she lacks strongly voluntary control. If strong doxastic voluntariness has anything to do with an agent's credence formation or her representation of reality, it is only as a tool with which she can promote – or prevent – the rational management of her representation of reality. Similarly, having oils and a paintbrush does not make one's paintings great.

Let's return to the first objection, that there are situations in which human agents do seem to exercise strongly voluntary choice in their credence formation. If this is so, then van Fraassen's voluntarism might still be true. He could argue that, even if strong voluntarism is not a component of rationality across all possible worlds, there might be some features of rationality which are possible-world dependent. Thus even if, say, coherence applies necessarily (as I take him to hold), there might be 'add-on' features governing the rationality of this or that type of agent in this or that epistemic context. Thus where agents do have strongly voluntary doxastic control, as they do actually (or so he seems to think), voluntarist norms apply.

What are we to make of this suggestion? There might be something to the thought that rationality has some context-dependent features. After all, it is predicated of doxastic agents, and there are more possible types of doxastic agent than can be counted. It would be odd to think that rationality looked exactly the same for all agents in all possible worlds. Recall for example Chapter 3's suggestion that there are rational constraints on which proposition a given experience might deliver up as evidence, or on the credences an agent may form in her evidence proposition. It is not at all clear that these constraints will be the same from one possible world and one agent to another. Thus in a world in which agents do have strongly voluntary control over their credence formation, voluntarism might well apply. Van Fraassen could fall back on the descriptive claim that human agents have strongly voluntary control, from which he could argue that ideal rationality *for human agents* includes not only the voluntarist constraints on their doxastic attitudes described above, but also strongly voluntary control.

However, I will argue that human agents do not have strongly voluntary control over their credence formation. Even if rationality is partly world-dependent, and even if voluntarist norms apply in some worlds, the actual world is not one of them.

Complex Evidence Cases

In support of the claim that some human updating is strongly voluntary, one might appeal to cases in which agents really seem to deliberate and choose between one credence and another. Such cases, which we'll call *complex evidence cases*, are characterized by an agent's acquiring a large or complicated quantity of evidence, much of which may be of a highly theoretical (as opposed to merely sensory) nature, and which does not at all clearly 'point', for that agent, toward one conclusion or another. For example:

> You are the only juror on a criminal case. Your task is not to pronounce 'guilty' or 'not guilty', but instead to provide a subjective probability of guilt, to the second decimal place.[140] You do not need perfect introspective access to your credences, because the judge has a credometer which perfectly tracks agents' credences to the second decimal place. The evidence presented is complex: the trial consists of many days of testimony by witnesses, experts, friends, and family members, on such diverse matters as the defendant's character, living circumstances, whereabouts at such-and-such a time, history, habits, and circle of acquaintances. DNA tests are offered but their provenance and reliability seem questionable (as far as you can tell, though you are aware that you are no expert at judging the opinions of scientific experts). This is to say nothing of the lifetime of background evidence you bring to the courtroom, in light of which you evaluate the evidence gleaned in your experience of the trial, which includes but is by no means limited to the material presented by the lawyers.
>
> When you return home after hearing each side make its case, planning to have your credence clocked by the credometer in the morning, you are torn. There are so many considerations, many of which are mutually undermining. If you evaluate the total evidence in a way which employs a simplicity criterion as the main epistemic desideratum, a credence of *x* seems unavoidable. But if you evaluate it in another way, which lends weight instead to explanatory breadth, a credence of *y*, which is different from *x*, seems equally unavoidable. Turning your attention to one or the other evaluation puts you in a temporary state which resembles having the credence in question, except that, when you turn your attention to the other evaluation, your state changes. Note that your total evidence includes propositions about your own wavering – this evidence factors into the respective evaluations resulting in *x* and *y*, respectively.
>
> Finally, you come down in favor of a credence of *x* in Not-Guilty, which the credometer duly records in the morning.

[140] The mechanics of the fine-grained justice system might be worked out like this: If your probability of guilt is greater than half, then the prisoner will be punished in a way that correlates with the value of your probability assignment. If your probability is less than half, then the plaintiff will be liable for the defendant's legal fees in a way which correlates directly with your probability assignment. If your probability is exactly half, there is a retrial.

The proponent of Some Strongly appeals to complex evidence cases such as this to motivate Some Strongly, noting that there is a phenomenology of turning evidence over in one's head, looking at it from different angles, rehearsing one's memories of acquiring it, and so forth, and that agents who undergo such phenomenology can and sometimes do wind up with one credence rather than continually wavering. What, if not the making of a strongly voluntary choice of credence, is going on here? The strong voluntarist concludes on the basis of this intuitive data that Some Strongly is true.

But I don't think that Some Strongly offers the best account of complex evidence cases. I propose the following involuntarist account: in such cases the agent is looking for more evidence. In particular, he is trying to introduce evidence which he does not yet grasp, in the form of a mental experience (see Chapter 3), to the effect that:

> *This* is the way to make sense of these apparently incompatible pieces of evidence (*Deictic Proposition*),

where 'this' refers to some theory which has not occurred to the agent before. The experience in which he came to grasp a proposition such as Deictic Proposition will, he hopes, generate credences in Not-Guilty conditional on Deictic Proposition, which may break the tie between credences of *x* and of *y* in Not-Guilty. He is, in other words, looking for more evidence to compel his credences in either direction, no matter which.

This involuntarist account is supported by common sense. Imagine that you are the juror in the complex evidence case and have gone home after the case closes, anticipating being scanned by the credometer in the morning. You are visited at home by a mafioso, who credibly threatens you with dire consequences if you do not form a specific credence in Not-Guilty, rather than suspending judgment between them. You have a strong motivation to meet his demand; how would you do so? You would not, I take it, try to change your credence merely by issuing commands to yourself. You are well aware that this will get you nowhere. You would introduce some indirect influence on your doxastic state, such as drugs or hypnosis; or, if these are not available, as they are not in our thought experiment, you would try to introduce evidence. This stands to reason, for the right evidence can guarantee that you form the credence you desire – and nothing else short of high precision chemical intervention can. The reason is that updating is not strongly voluntary.

Conclusion

Because of certain liberationist objections to Bayesianism which will be discussed in the next two chapters, van Fraassen rejects Bayesian Updating, replacing it with voluntarism. Voluntarism is non-Evidentialist in that it allows the agent's will to be another factor, beside her evidence, which rationally permits or mandates updates; it is more specifically non-Evidentialist in the Bayesian sense, since it denies that an agent must update by conditionalizing on her evidence. I have argued that voluntarism is untenable, on the grounds that the notion of commitment is more trouble than it is worth, that it succumbs to a form of subjectivism about rationality, and that strongly voluntary updating has nothing to do with ideal rationality, even a notion of ideal rationality governing human agents.

However, this chapter has not looked at van Fraassen's reasons for rejecting Bayesianism to begin with. If they are on target, then this would be unfortunate indeed, since voluntarism cannot stand as an alternative. However, Chapters 6 and 7 will show that they are not on target. In airing them van Fraassen has put his finger on some important worries about a common Bayesian package. However, as Chapters 6 and 7 show, it turns out that van Fraassen's major criticisms are met by Expansive Bayesianism, the view I am developing, which keeps Bayesian Updating (and hence Unique Rule and Rulebound), and instead rejects some extraneous assumptions which have grown up around the Bayesian foundation.

5. Bayesian Updating and Doxastically Impossible Evidence

Introduction

This chapter introduces some key elements of Expansive Bayesianism, by showing a commonly assumed Bayesian claim to be false. The claim is:

> Once a rational agent's credence in a proposition *p* has reached 1 (and hence her credence in not-*p* has reached 0), any change of credence in *p* (or in not-*p*) is irrational. (*Unrevisable*)

The following argument is given for Unrevisable:

> *Argument from Conditionalization*
>
> 1. Conditionalization is the only rational way for an agent update her credences. (Bayesian Updating)
>
> 2. No conditionalizing update can move a credence down from 1 or up from 0. (*Stuck*)
>
> 3. Hence Unrevisable is true. (from 1, 2)

I will argue that Unrevisable is false.[141] The Argument from Conditionalization is unsound, since Premise 2, the claim I call 'Stuck', is false. But before addressing the rationality of updates down

[141] The question whether Unrevisable is a desirable feature of rationality (as opposed to merely entailed by Bayesianism) is discussed by Skyrms (1987), Levi (1967), and Williamson (2000). Levi and Williamson each offer

from 1 or up from 0, I'll spend a section addressing their possibility: if they are not even possible, then the question whether they are rational is moot.

Is it possible for credences to move down from 1 or up from 0?

One might deny that receiving evidence can prompt an agent to change an extremal credence, asserting:

> Once an agent has assigned credence 1 to a proposition *p*, it is impossible for him to receive evidence which brings his credence in *p* down from 1 (and thus his credence in not-*p* up from zero). (*Impossibility*)

The modality of Impossibility is ambiguous; the relevant alternatives include conceptual and what I'll call 'nomological' impossibility. I'll show that Impossibility, interpreted as pertaining to either modality, is false.

First, conceptual modality:

> Once an agent has assigned credence 1 to a proposition *p*, it is *conceptually impossible* for him to receive evidence which brings his credence in *p* down from 1 (and thus his credence in not-*p* up from zero). (*Conceptual Impossibility*)

A proposition *p* is *conceptually impossible* just in case, by rules of linguistic meaning, a contradiction can be derived from *p*. The terms the meanings of which would most plausibly substantiate Conceptual Impossibility are 'evidence' and 'credence'. The thought would be that the definition of either term makes it self-contradictory to suggest that there could be evidence which brings an agent's credence down from one or up from zero.

Conceptual Impossibility is false, at least according to the ways in which I am using the terms in question, which I have defended in previous chapters. To recall from Chapter 3: The evidence an agent receives (which Bayesians call her 'new evidence') is a proposition which an evidential experience presents to her and in which the agent invests a new credence. This conception of evidence does not prohibit credences from moving up from zero or down from 1. And recall also from Chapter 3 that a credence can only be had toward a proposition which the agent grasps. This entails nothing about the circumstances in which a credence can change its value, or about what values it may permissibly have. Hence in order to defend Conceptual Impossibility, one would have to argue in favor of different conceptions of 'credence' and 'evidence' than I employ.

---

non-Bayesian accounts of updating which are motivated in part by the assumption that Unrevisable does follow from Bayesianism.

In this vein, however, some Bayesians might define credence 1 in *p* as the disposition to ignore all evidence against *p*. But this move will not work. First, even if credences are defined in this way, Conceptual Impossibility does not follow. Additional arguments would be needed to support the assumption that dispositions cannot be lost or overriden (say, by the agent's having a vivid experience as of *p*). The friend of Conceptual Impossibility may respond that *some* dispositions can be lost but that others, in particular the disposition to ignore evidence against a proposition of which one is certain, cannot. But this move appears *ad hoc*, at least without a reason to think that the disposition to ignore evidence against *p* is among the dispositions which cannot be lost. Hence, short of such a view being spelled out in a way that addresses these problems, we will not worry about Conceptual Impossibility.

A more plausible version of Impossibility pertains to the modality of what can and cannot occur given the properties of the actual universe:

> Once an agent has assigned credence 1 to a proposition *p*, it is *nomologically impossible* for him to receive evidence which brings his credence in *p* down from 1 (and thus his credence in not-*p* up from zero). (*Nomological Impossibility*)

By the claim that *p* is nomologically impossible, I mean that the laws which obtain in our universe[142] ensure that *p* is false. This is more than the claim that *p* will not actually obtain; it is the claim that, because of the properties of our universe, *p cannot* actually obtain. Features of the universe which might make Nomological Impossibility true, the thought is, include such things as the cognitive properties of human agents or the properties of credences.

Perhaps the most plausible thought in this vicinity is that having a prior credence of 0 in *p* makes a human agent incapable of forming a non-zero credence in *p*, no matter what he experiences. Put another way: his credences restrict the type of experience which he is capable of undergoing. They may even make some experiences impossible for him, namely, those which might tend otherwise to suggest (to agents who already have some non-zero credence or other in *p*) a particular non-zero value for *p*. According to this thought, then, an agent's prior credences are at least sometimes capable of *determining* what evidence that agent can receive from any given experience.

The suggestion is plausible that credences have some influence on what agents experience and thus on their evidence, and is often discussed under the heading 'the theory-ladenness of

[142] The proponent of Nomological Impossibility need not be committed to realism about laws of nature; such laws might be descriptions of properties of the universe, rather than rules which govern in some more robust sense.

observation'.[143] For example, someone who is certain that he is unforgiven or unloved might face great obstacles to experiencing forgiveness or love; he will tend to interpret his experiences of his own feelings, or of the gestures of others, in ways that do not at all raise his subjective probability for such propositions as *I am forgiven* or *I am loved*. But to say that he is *incapable* of interpreting his experiences of such things – indeed, of anything – as raising his probability for these propositions is a strong thesis indeed. It is to say that, no matter what kindness, affirmation, self-sacrifice, or encouragement are directed his way, he will – indeed, he is predetermined to – always remain certain either that they are not really kindness, affirmation, self-sacrifice, or encouragement, or that they are not indications of love or forgiveness.

Perhaps such cases arise sometimes. But Nomological Impossibility says more: it says that, with actual agents, they occur *necessarily*; indeed, *whenever* an agent has a credence of 0 or 1, because the universe is so constituted. This is an exceptionally strong thesis, much stronger than the plausible claim that experiences are affected by prior credence (or 'laden by theory'). We may call it the claim that experiences are *saturated* by prior credence:

> Agents who have zero credence that *p* are incapable of having either experiences as of *p* or experiences which raise *p*'s subjective probability above zero. (*Saturation*)

Another way to put Saturation might be: such agents are incapable of receiving evidence favoring *p* which is not ruled out by their prior credences. Saturation entails Nomological Impossibility. Nomological Impossibility does not entail Saturation – moving from a zero to a non-zero credence can be nomologically impossible for other reasons – but I take it that Saturation is the most plausible way of supporting Nomological Impossibility.

Saturation is a terribly strong claim about the properties of experience and of credence. Why think that it is true? One might be given by the observation that agents exhibit such behavior: many agents simply cannot, it seems, lower their credence in a proposition of which they are certain. Take for example the mother who cannot be brought to lower her credence, even by what would, for anyone with a less than certain credence, be condemning evidence, in the proposition *My son is innocent of the murder*. Fingerprints, testimony of hitherto thoroughly trusted sources, videos of the crime, even a frank and clearly uncoerced confession from her son to her own face in the presence of no other witnesses, will do nothing to sway her utter confidence in his innocence. She interprets all of this evidence not as disconfirming *My son is innocent of the murder* (for this is certainly true, by her lights), but instead as confirming some other proposition, such as *I am dreaming*, *A conspiracy*

[143] For original statements of strong views about the theory-ladenness of observation, see Popper (1963), Lakatos (1978), or a somewhat more subtle view, Sellars (1956). For helpful discussions of theory-ladenness see Brewer and Lambert (2001) and Kosso (1992), Chapter 6.

has been unleashed against me and my family, *The alleged evidence for my son's guilt has been cunningly fabricated by aliens*, or something along these lines. As far out as such cases might seem when apparent counterevidence builds up, I won't deny that they are possible. But that such cases are possible does not support Saturated, which is the claim that *all* cases of certainty in *p* make it impossible for agents to become less than certain in *p*.

Surely Saturated is false. Counterexamples abound. Imagine for example a sharp undergraduate who (is weakly coherent and hence) assigns credence 1 to (P v ~P). He rationally ought not change his credence, since (P v ~P) is a logical truth and must rationally receive credence 1 if it receives any credence, but surely he is *capable* of changing it if he does invest credence 1 in it. Imagine that he undertakes graduate work in philosophical logic and becomes confident that a non-classical logic is true; this lowers his credence in (P v ~P). But if Saturation is true, he could not lower his confidence in it. If Saturation is true, it would follow that the only agents who could become confident that non-classical logics are true are those who *start off* incoherent – who were never certain of logical truths to begin with! For if they had been certain, they would be incapable of having any non-zero credence, to say nothing of outright confidence, in a non-classical logic.

The proponent of Nomological Impossibility may respond that agents such as our student must be confused about their credences. After all, credences are surely not luminous (i.e., surely not such that an agent can always introspect that his credence is *x* just in case his credence is *x*). The case of the sharp undergraduate is in fact nomologically impossible. What might look to external observers or to the agent himself like a case of certainty in (P v ~P) is really one of two things. Either it is just a case of very high confidence; or it is a case of *staying* certain that (P v ~P) and being mistaken about his non-zero credence in classical logic. That is, the supporter of Saturated *defines* certainty in such a way that Nomological Impossibility turns out true. Credence 1 just is certainty so absolute that agents are incapable of receiving evidence to thwart it.

How plausible is this suggestion about certainty? One problem is that it seems to suggest that agents are often and chronically, rather than exceptionally, seriously mistaken about their own credences. This suggestion, I take it, is a desperate leap to save an implausible principle. Moreover, what use does an epistemology have for such a notion of certainty? It may track an interesting psychological state which appears in situations of extreme dogmatism, as in the mother case above. But such cases are surely a rare exception. I conclude that Saturation, and hence Nomological Impossibility, are false: it is possible, given the laws by which the universe operates, for agents to move from certainty to uncertainty. But it is another question entirely whether such a move is rational. This is just what proponents of Unrevisable deny. I'll introduce some key features of Bayesianism for the discussion to follow, and then I'll argue that Unrevisable is false.

Rational Credence Change in Bayesianism

A typical Bayesian take on Unrevisable is given by Meacham, who says that conditionalization "requires certainties to be permanent: once you're certain of something, you should always be certain of it" (2010, p. 1). The reason, he says elsewhere, is that "you can only lose doxastic [possibilities] in this process, not gain them. Being certain of a proposition P entails that all of your current doxastic [possibilities] are compatible with P, and if you only lose doxastic [possibilities] when you update, then all of your future doxastic [possibilities] will be compatible with P as well" (2008, p. 3).

This view exemplifies the pictorial model of updating which van Fraassen calls the 'muddy Venn diagram' (1989, Chapter 13). Imagine a Venn diagram of which each cell represents a set of doxastic possibilities. When a cell has mud on it, the possibilities within it are live for the agent, and when the mud is removed, so are the possibilities. On this model, becoming certain of an evidence proposition E is represented by removing any mud from the cell representing not-E, and assigning E an uncertain credence is represented by the shifting of some but not all of the mud from E to not-E or vice versa. Conditionalization is represented by the redistribution of the removed mud over the remaining cells, in proportion to the amount of mud which those cells already had. On the muddy Venn model evidence can do only one of two things, depending on whether the agent assigns it a credence of 1 or less than 1. If she assigns it a credence of 1, then her space of doxastic possibilities *shrinks*: possibilities inconsistent with her evidence are deleted from the diagram when all of the mud is removed from them; there can never be mud on them again. (The removed mud is then redistributed over the remaining cells in proportion to the amount of mud they had already.) If the agent forms a credence of less than 1 in her evidence, then the mud is redistributed among all of the cells which still have mud on them.

Thus on the muddy Venn model of updating, the agent's posterior credence distribution is always a sub-model of the credence distribution which she has before assigning any credence to her evidence. If this model is taken to be normative over updating, then there is no room for the agent either to 'recover' deleted cells or to 'create' cells where there were none before.

But of course, a picture is only a good heuristic if it is true to reality in the relevant ways. If in fact agents *can* rationally invest non-zero credences in evidence in which their credence was not previously greater than zero, then the muddy Venn picture of ideal credence formation misses its mark. It becomes a hindrance to understanding, and the picture must be discarded or adjusted. The upshot of the rest of this chapter will be that this picture needs adjustment.

I'll outline two ways in which credences can rationally change in Bayesianism, and will show that the muddy Venn model of updating builds in more strict requirements than Bayesian Updating itself does.

Two Ways in which Credences Can Rationally Change

This section outlines the two ways in which, according to Bayesianism, credences can rationally change. One way excludes the possibility that a credence of 1 can rationally become less than 1 (or a credence of 0 greater than 0). The other is commonly taken to exclude this as well, but this chapter argues that in fact it does not do so: Unrevisable is false.

Bayesianism has it that a credence can rationally come about in only three ways. First, it can be an *initial* credence, which is to say, a credence with which the agent simply finds herself and has not had occasion to change in light of any evidence. An agent's initial credence in *p* is her credence in *p* conditional only on the tautology. Second, a credence can be the result of updating a previous, or an initial, credence by conditionalization on evidence; in this case her credence in *p* is her credence in *p* conditional on the total evidence she has thus far received (as well as on the tautology). Third, a credence can result from a representational experience: the agent has an evidential experience which presents her with a proposition; she initiates the update by investing a credence in that proposition.

When a credence is greater than 0 or less than 1, there are two ways in which it can change its value: either by being conditionalized on new evidence, or by the agent's investing a credence in a proposition which an experience has generated as evidence. Extremal credences (i.e., credences of 1 or 0) are a special case, however. They cannot change in the first way, by being conditionalized *on* any new evidence; Bayes's Theorem excludes such an eventuality:

$$\mathrm{Cr}(P|E) = \frac{\mathrm{Cr}(E)\mathrm{Cr}(E|P)}{\mathrm{Cr}(E)}, \qquad \text{where } \mathrm{Cr}(E) > 0$$

(*Bayes's Theorem*)

The reason is, when the value of H is 1, then Cr(H|E) will always come out 1 after conditionalization. And when the value of H is 0, Cr(H|E) will always come out 0 after conditionalization.[144]

[144] This is true both of straight and of Jeffrey conditionalization (see Chapter 2). Straight conditionalization proceeds by a straightforward application of Bayes's Theorem. The only difference in Jeffrey conditionalization is that Bayes's Theorem will be applied to each member of a logical partition; and since a partition has a total probability of 1, when Cr(H) = 1, the value of Cr(H|E) when E is less than 1 will always remain 1; and when Cr(H) = 0, the value of Cr(H|E) when E is less than 1 will always remain 0.

This leaves one possible way for extremal credences rationally to change: by receipt of evidence.[145] An agent might have an experience which presents to her a proposition which was, before the experience, doxastically impossible for her; this amounts to having a credence of 0 in *p* yet having an experience change it to a value greater than 0, or to having a credence of 1 in *p* and having an experience change it to a value of less than 1. We'll call such situations *impossible-evidence scenarios* (where the modality of 'impossible' is doxastic, and the evidence was impossible until the evidential experience). A proposition *p* is *doxastically possible* for an agent at a time *t* just in case the agent, at *t*, assigns a non-zero credence to *p*. What I'm calling 'impossible evidence' for an agent is a proposition in which that agent, at time *t*, has zero credence, and to which, on having some evidential experience at *t+1*, she then (at *t+1*) assigns a non-zero credence.

Proponents of Unrevisable deny that impossible-evidence scenarios are ever rational. The Argument from Conditionalization holds them to be irrational because it assumes that the caveat in Bayes's Theorem, 'where Cr(E)>0', excludes impossible evidence from being conditionalized on.

I've already argued that such experiences are nomologically possible; next, I'll argue, against the proponent of premise 2 of the Argument from Conditionalization, that they are rational.

The Argument from Conditionalization

Let us look more closely at the Argument from Conditionalization, which concludes with Unrevisable:

Argument from Conditionalization

1. Conditionalization is the only rational way for an agent to update her credences. (Bayesian Updating)

2. No conditionalizing update can move a credence down from 1 or up from 0. (Stuck)

3. Hence Unrevisable is true. (from 1, 2)

We will leave premise 1 alone, since it just states the Bayesian view of rational updating. The idea is to see whether this view is compatible with Unrevisable, which premise 2, Stuck, denies. But Stuck is false, as we will see; hence the Bayesian view of rational updating is compatible with Unrevisable.

[145] There is an additional way in which credences can change which Bayesianism does not consider rational yet which bears mentioning: forgetting. Forgetting results neither from conditionalizing on evidence nor from evidential experiences, but merely from the passage of time. Yet it certainly rivals both of these in terms of how often it occurs. If forgetting is modelled as in Williamson (2000), namely as the agent's credence in *p* gradually changing as time passes, then forgetting is on my view irrational since it violates Bayesian Updating.

We saw above that extremal credences cannot change by *being* conditionalized *on* evidence, but we also saw that it is possible, both conceptually and nomologically, for an extremal credence to change if an agent receives *as* evidence the very proposition which is its object and invests a non-zero credence in it. If Stuck true, then, even though the receipt of evidence which was hitherto doxastically impossible is nomologically possible, such evidence cannot be conditionalized on and hence (since there is, we are assuming, no other rational way to update) cannot rationally move the agent's credence to a non-extremal value. But why would one think that Stuck is true, and thus that impossible evidence cannot be conditionalized on?

To see why one might think this, let us look more closely at an updating scenario, which is divisible into two stages: one prior and another posterior to updating.

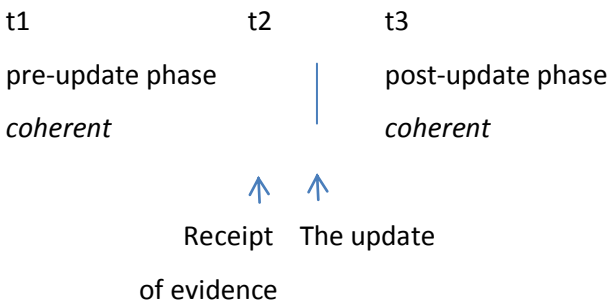

Figure 1: The Updating Scenario

At time *t1*, the agent is in the prior phase, not having received new evidence since her most recent update (whenever that was). At later time *t2*, she has an evidential experience, which leaves her with new evidence on which to update.[146] Next, the agent updates. Since the update may (as far as I can see) be simultaneous to the receipt of evidence, I have not allotted it a discrete time point on the diagram. The next time is *t3*, when the agent has updated and is in a position for the cycle to begin again the next time she receives new evidence.

Because the update is rational, it is a conditionalizing update. Recall from Chapter 2 that

> A *conditionalizing update* is one in which the agent sets her new credences equal to the value of her old credences conditional on her new evidence. (*Equal*)

[146] An agent can assign a credence to her evidence which does not mandate conditionalization since she has no credences conditional on it, but I'll ignore such cases for now.

Thus the rational agent's posterior credences take on the same value as her prior conditional credences – which is to say, the same value as her prior credences conditional on her newly received evidence.[147]

Now we can explain the temptation for Bayesians to endorse Stuck, premise 2 of the Argument from Conditionalization. Note that Bayes's Theorem (given above), which states how to derive a posterior credence from a prior credence plus the agent's credence in her evidence, has a caveat: it can only be applied when the agent's prior credence in her evidence is greater than 0. At first glance this might seem to rule out impossible-evidence scenarios, since the agent's prior credence in the evidence in such cases is 0. The thought, I take it, is this: in order for the agent to conditionalize her prior credences on any evidence, she must have prior credences conditional on that evidence. But when her prior credence in an evidential proposition E is 0, she cannot have prior credences conditional on E. The reason is, Bayes's Theorem requires a non-zero denominator; hence plugging a zero into the evidence place in the denominator results in an undefined conditional credence. Any prior credences conditional on the impossible evidence would be undefined. Hence the thought that impossible evidence can never be updated on by conditionalization.

But taking this argument more slowly will show that its conclusion is false: sometimes, impossible evidence *can* be updated on by conditionalization:

Argument for Stuck: Impossible-evidence scenarios preclude conditionalization

(A) In order to update a prior credence in H on evidence E by conditionalizing, the agent must, in the pre-update phase, have a credence in H conditional on E on which to conditionalize. (by definition of conditionalization)

(B) An agent who has a zero credence in E cannot, before updating, have any credences conditional on E. (Assumption)

(C) Therefore, such an agent cannot, on receipt of impossible evidence E, update by conditionalizing on E. (from 1, 2)

This argument is valid but it is unsound. Premise B is the culprit, and indeed three times over; there are three models of possible updating scenarios which falsify it: (1) a model in which the agent already has credences conditional on the impossible evidence, even though he does not have non-zero unconditional credences in that evidence; (2) a model in which the agent acquires conditional credences at the same time as he assigns the non-zero credence to his evidence; (3) a model in

[147] Jeffrey conditionalization requires additional steps, but the points I am making here do not change for Jeffrey conditionalization. For simplicity I'll stick with an exposition in terms of straight conditionalization, that is, conditionalizing on evidence which has newly acquired credence 1.

which the agent acquires a credence in a proposition which he did not previously grasp and which provides him with evidence *for* his impossible evidence which he could not have imagined receiving.

(1) Conditional Credences as Primitive

On our first model, in the pre-update phase, before having the evidential experience which gives rise to impossible evidence E, the agent *already has* credences conditional on E. How can this be, if her credence in E in that phase is zero? Aren't conditional credences just defined in terms of unconditional credences, so that no credence conditional on a doxastically impossible proposition can be defined? This claim is widely assumed among Bayesians:[148]

> An agent's credence in H conditional on E (for short: H|E), at any time *t*, is equal to the value derived by plugging her unconditional credences at *t* in H and in E into Bayes's Theorem. (*Conditional Reduction*)

I call this claim 'Conditional Reduction' because it asserts that conditional credences at a time are reductively defined in terms of unconditional credences at that time. If it is true, then, whenever an agent lacks unconditional credences, she thereby also lacks conditional credences, and hence she cannot conditionalize. If Conditional Reduction is true, then our agent cannot have credences conditional on E in the pre-update phase, since her credence in E itself is zero; and hence, if the evidential experience fails to deliver up credences conditional on the hitherto impossible evidence, she cannot conditionalize on that evidence.

But Conditional Reduction is false, which Hájek argues (to my mind, conclusively). He provides a litany of examples in which an agent clearly has conditional credences – and rationally should have them – in spite of having either no unconditional credences or having unconditional credences of zero.

One of Hájek's examples is continuous random variables: The probability that a continuous random variable X takes on a particular value x equals 0. Even so, it appears that various probabilities (and hence credences, on the assumption that credences ought to equal any certain chances that there may be), conditional on the proposition that X equals x, are well defined, such as:

> Cr (X = x, given X = x) = 1
>
> Cr (X ≠ x, given X = x) = 0
>
> Cr (T, given X = x) = 1, where T is the tautology

[148] For discussion see Háyek (2003).

Cr (*this fair coin lands heads*, given X = x) = 1/2

On an intuitive level this suggestion makes sense. Here is another example: You are certain that the moon is not made of green cheese. Nonetheless, applying a principle of statistical inference to the effect that, if one celestial body is made of some substance, then this raises the probability that others are made of it too, you have credences conditional on the impossible proposition *The moon is made of green cheese*. For example, you have credences in *Asteroid 1 is made of green cheese*, *the Sun is made of green cheese*, and so forth, which are higher given *The moon is made of green cheese* than given its negation. We can still imagine what the world would be like if doxastic impossibilities were true and we can still reason about what other propositions those impossible ones make more likely.

But if conditional credences are not defined in terms of unconditional credences, what are they? Hájek makes the alternative suggestion that the order of definition should be reversed: that conditional credences be taken as basic instead, and that unconditional credences be defined in terms of them. That is, he suggests replacing Conditional Reduction with:

> An agent's unconditional credence in H, at any time t, is equal to the value of the agent's credence in H conditional – in accord with Bayes's Theorem – on the tautology. (*Unconditional Reduction*)[149]

This is a sensible suggestion, and I endorse it. Hence having conditional credences which are defined even though the agent's credence in the evidence is 0 is one way to falsify premise (B) of the argument that impossible-evidence scenarios preclude conditionalization: an agent can have a non-zero credence in H conditional on E which remains dormant until an experience changes the agent's credence in E from zero to some positive value.

(2) Receiving Impossible Evidence

But for those who still endorse Conditional Reduction, there are two more models which falsify premise (B). These models become clear when we consider the update itself in greater detail. It proceeds in two steps, which are logically if not temporally distinct. The first, which we may call the

[149] Note that there is still a connection between conditional and unconditional credences:
Cr(H|E)Cr(E) = Cr(H^E). (*Link*)
Since Cr(E) is not 1, Link in turn entails that Cr(H^E) is not equal to Cr(H|E). In impossible-evidence scenarios, for impossible evidence E, Cr(H^E) = 0; but this is clearly consistent with the conditional credence taking any probability value at all. Of course, when Cr(E)>0, Link does entail that the agent's conditional credence Cr(H|E) equals her unconditional credences in H and in E plugged into Bayes's Theorem.

input step, occurs when the agent assigns a credence to her newly acquired evidence. The second step is conditionalization on the evidence, using the value which the agent assigns to the evidence in the input step. We may thus fill in the above diagram as follows:

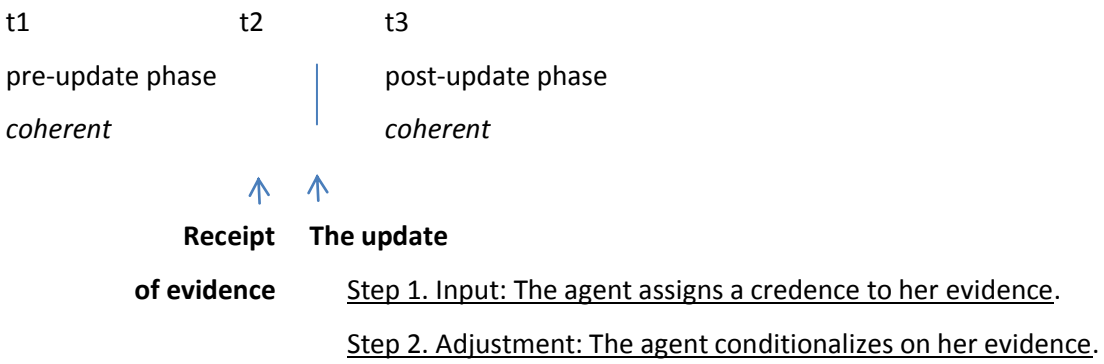

Figure 2: The two steps of the update

Distinguishing these two steps of the update reveals another way in which premise B can be falsified. Let's zero in on step 1 of the update, the inputting of a new value – or the assigning of a new credence – to the agent's evidence.

I have argued above that it is nomologically possible for an agent to assign a non-zero credence to an evidential proposition in which, before she had the evidential experience, she had zero credence. I submit that doing so can be rational. If an agent both can and can rationally assign a non-zero credence to any evidence proposition at all, no matter what its previous value, then there seems every reason to think that she can *also* assign values to *completely new* conditional credences. Before the evidential experience, the agent did not have credences conditional on the newly possible proposition, since it was previously doxastically impossible. But now it has become doxastically possible; and, if it is possible, then it stands to reason that other things must be possible given it. There are gaps in the agent's credence distribution which must be filled in, and this must occur by the assigning of completely new conditional credences. Thus the agent's post-update credence distribution may differ from her pre-update credence distribution by containing an updated credence in her evidence proposition E, *as well as* by containing completely new credences conditional on E which her pre-update credence distribution did not contain. If the agent does assign new credences to various other propositions given the previously impossible E, then – since she does so prior to any update on E – she can just conditionalize in accord with the new conditional credences. Thus the agent's pre-update credence in E is non-zero, so she can conditionalize on E without violating the caveat, as long as she acquires credences conditional on E as well.

To this suggestion, one might object that it is utterly mysterious how an agent can simply generate new conditional credences. It is bizarre to suppose that conditional credences just pop into existence the instant they are needed – the suggestion has the slimy feel of a *deus ex machina*. Perhaps it is logically possible that such an eventuality occur, but actuality gives no more reason to think that conditional credences can pop into being to explain irrational updates than that fairies and gnomes should be posited to explain the disappearance of socks from the laundry.

This objection is wrong-headed. The question of where the new conditional credences come from is a specification of the question of where *any* initial credences come from. This is one of the major mysteries of Bayesianism;[150] I challenge the objector to solve it, or else to give up the whole idea of initial credences and hence Bayesianism itself. Even so, the fuller a story I can give about the origin of the new conditional credences, the more plausible my suggestion will be (although I don't find it implausible even without some such story).

One idea is that they are generated by the same experience which delivered up the impossible evidence. This stands to reason; experience in one fashion or another is likely the root, either as a determining or a contributing cause, of all of an agent's credences, and the impossible-evidence experience may well be just one more of those. Another idea is that the agent, on having the impossible-evidence experience, pauses to reflect on how the new, non-zero credence in E affects the probability of the other propositions over which her credences are defined, and the new conditional credences are the fruits of these reflections.[151] There is no reason to suppose that update cannot be temporally extended in such a fashion.

In short, the question where the new conditional credences come from is not much of a challenge to my suggestion that Stuck can be falsified by scenarios of the type described. And anyway, even if this question remained more mysterious than I have shown it to be (or at least, more mysterious than the question of where credences come from to begin with), the logical possibility of my suggestion is enough to show premise (B) to be false.

Indeed, the claim called 'Equal' bears repeating: conditionalization is merely the process of setting one's credences equal to the conditional credences which one had before conditionalizing. There is nothing about conditionalization which requires that the conditional credences be present *before* the agent invests a credence in her evidence. All that is needed to satisfy Bayes's Theorem is that the agent have a non-zero credence in her evidence, and credences conditional on that evidence, before the update. These are just logical prerequisites for the update itself to occur. Hence

[150] For a suggestion about where priors come from see Suppes (2007).

[151] Jeffrey (2004) suggests that "[y]our 'subjective' probability is not something fetched out of the sky on a whim; it is what your actual judgment *should* be, in view of your information to date and of your sense of other people's information' (p. 76).

premise (B) is false: an agent who has zero credence in her evidence, E, can, before updating, have credences conditional on E. Hence the conclusion, which is Premise 2 of the Argument from Conditionalization (Stuck), is false: agents can update on impossible evidence by conditionalizing. This is so even if, pace Háyek, unconditional credences are reductively defined in terms of conditional credences. But it may be so if instead conditional credences are basic. The models discussed in (1) and (2) may both be true.

(3) Receiving Undefined Evidence

When we take on board the possibility that agents can acquire new initial credences, we see that there is yet a third model which falsifies premise (B). This model, like the previous one, does not depend on the claim that unconditional credences are basic. What it does depend on are the claims defended in Chapter 3, namely, Grasping, that credences can only be had in propositions that one grasps, and Newly Grasped, that a newly grasped proposition can be evidence.

The idea is this: In the event that an agent comes to grasp a new proposition, she may also assign a credence to that proposition. In doing so she acquires a completely new credence, in the following sense: Before the agent invested a credence in the newly grasped proposition *p*, her credence distribution C1 was defined over some number *n* propositions; after she invests a credence in *p*, C1 is replaced with a credence distribution C2, defined over more propositions (that is, over *p* and any of its truth-functional combinations in which the agent also has a credence). Coming to grasp *p* has the effect of broadening the agent's doxastic possibility space to include another way which the world could be, which she had not entertained before coming to grasp *p*. It is what Chapter 1 called a 'C-set expansion': an expansion of the set of propositions over which the agent's credences are defined. Hence coming to grasp *p* is an evidential experience like any other, except that its representational content is completely new to the agent. The shift from C1 to C2 is the same kind of shift that occurs when an agent acquires evidence which was defined before the evidential experience. We'll call evidence E that arises in this way *undefined* evidence, by which I mean, evidence in which the agent's credence was undefined before she has her evidential experience.[152] The claim 'Newly Grasped' establishes that experiences can provide evidence in which the agent's credence was undefined before she has the experience.

[152] There is a second way in which agents can receive undefined evidence: by receiving as evidence a proposition in which they had, before the experience, suspended judgment and hence which was previously undefined for them. One might think that withholding judgment is irrational but nothing on my model says that it is. But I will not discuss here the receipt of undefined evidence in which one had previously suspended judgment.

But what does this have to do with impossible evidence, which is evidence which the agent could grasp before the experience which generates it as evidence for her? Undefined evidence can be evidence *for* a doxastically impossible proposition. This stands to reason; if anything would bring an agent to invest a non-zero credence in a proposition which was previously impossible for her, it would be the cognizing of utterly new doxastic possibilities.

Let us consider a few examples.

Examples of Impossible-Evidence Conversions

An investigative journalist has a high prior credence in the proposition

G People are basically good,

even conditional on the proposition

E Mass murders are committed by drug cartels.

In one scenario, the journalist acquires a high credence in E by reading a newspaper. Conditionalizing on her total evidence, which is the conjunction of E and her background evidence B, lowers her credence in G a little but not significantly. But in another scenario, she acquires a high credence in E by doing on-site reporting and unearthing a pit of bodies on a ranch in the desert. Here, of course, she has not just acquired a high credence in E, but also in a host of other propositions which become part of her total evidence, which we'll call (E ^ B*). One of the propositions in B*, which was not in B, is

D Mass murder has results like that,

where 'that' encapsulates the gruesome phenomenology of uncovering the bodies and thereby taking herself to be confronted with human evil in concrete way.

Before the journalist had the experience, her credence distribution did not include D at all, because she could not have conceived of the experience to which 'that' refers. D is either an entirely new possibility which enters her doxastic space, or a proposition too fine-grained to have been transparent before her experience. Coming to grasp it and invest a non-zero credence in it provides her with evidence on which she could not, before the experience, have had conditional credences.

Consider another example of an agent who receives ungrasped evidence which provides unexpected yet significant support for a proposition which, before the experience, was doxastically impossible. Imagine that Ben has a conditional credence in

L I would give up my life willingly

which is quite low, almost zero, and his conditional credence in L given

E1 I become a father

is the same. That is, L is indifferent, for Ben, with respect to the evidence E1. But when his new baby is put in his arms his credence in L jumps up and stays up. He notices this by introspection, but it was the experience of the baby that brought about the change of conditional credence. In particular, the experience did not just bring him to become certain (or at least highly confident of) E1 but also brought him to become certain (or highly confident) that

D1 Having my own new baby in my arms is like that,

a proposition of which he could not previously conceive. Assigning a high credence to D1 brings him to change his conditional credence (L|E1). In particular he finds that his own life is worth less to him now than before learning D1, and he finds by introspection that many other credences, such as his credence in

L* I will not be content unless I become CEO,

which he had previously taken to be independent of E1 but now does not, are different too: in particular his certainty in L* has now become a certainty in not-L*, since his current state of contentment without being CEO flies in its face.

Thus, even though an agent does not have a credence in such evidence before she assigns a credence to it, as long as she acquires a credence in it by the time she updates, she can update on it by conditionalization. Hence undefined-evidence scenarios also falsify premise (B).

Conclusion

In summary, an agent may have defined credences conditional on some piece of evidence E before she updates on E. Thus I repeat that Stuck, Premise 2 of the Argument from Conditionalization, is false: agents can update on impossible evidence by conditionalizing. This means that the Argument from Conditionalization does not support Unrevisable, the claim that, once an agent's credence in a proposition *p* has reached 1 (and hence her credence in not-*p* has reached 0), any change of credence in *p* (or in not-*p*) is irrational.

This is the foundational insight of Expansive Bayesianism. In arguing against Unrevisable we have seen that Stuck, the claim that no conditionalizing update can bring a rational agent's credence down from 1 or up from 0, is false. Extremal credences can be changed by a conditionalizing update, in the event that the agent receives as evidence a proposition in which, before receiving it, his credence was either zero or not defined at all. Not only are such phenomena both conceptually possible and possible given the laws of the universe, they can also be rational.

Chapter 3 promised that Expansive Bayesianism would address the problem of new theories. Recall that this problem arises when a rational agent has credences summing to 1 in a list of rival theories, and comes to grasp and invest a credence in a totally new theory, H, which is not on that list. Surely, the idea is, she is rational to assign a credence to H. Bayesianism accounts for her doing so by stipulating that the rational agent has, for every list of theories, a non-zero credence in a 'reserve hypothesis' of the form *There is a theory which I have not yet grasped*, from which she can take some of her credence and assign it to the newly grasped theory. Stipulating that the agent have a non-zero credence in a reserve hypothesis amounts to requiring her not to be certain that the other theories exhaust the options. This is a fine solution, except that it faces a problem in cases where the agent happens, unfortunately, to be so dogmatic as to lack a reserve hypothesis. Such an agent will have credence 0 in *There is a theory which I have not yet grasped*. If the reserve-hypothesis solution is the only way to allow her update to a non-zero credence in the new theory to proceed by conditionalization, then one might worry that there is no rational way for this agent to assign a non-zero credence to the newly grasped theory.

Expansive Bayesianism shows this worry to be unfounded: the agent can be rational in assigning a non-zero credence to it after all. She can have an impossible-evidence experience which brings her to assign a non-zero credence to *There is a theory which I have not yet grasped*. A plausible candidate would be an experience in which she comes to grasp H itself; this might be a mental experience of coming to grasp it on thinking hard about it. Solving the problem of new theories thus does not require Bayesianism to stipulate that rational agents must have a non-zero credence in a

logically contingent propositions. In the event that the agent has an impossible-evidence experience, she can conditionalize even if, before the experience, she had credence 0 in *There is a theory which I have not yet grasped*. Resolving the problem of new theories is just one small way in which Expansive Bayesianism is a powerful and flexible form of Bayesianism.

Chapters 6 and 7 will continue in this vein, deploying Expansive Bayesianism to respond to some objections, launched in particular by van Fraassen, to the effect that Bayesianism is too strict to allow for certain types of rational credence change. Van Fraassen has done a great service to Bayesian epistemology in pointing out these problems, but they are not problems with Bayesianism itself, but rather with a form of Bayesianism which, either implicitly or explicitly, builds some additional assumptions onto the Bayesian framework in a way which curtails the power and elegance of that framework. The next chapter responds to the objection that a Bayesian theory of rational updating cannot model rational conversions.

6. Bayesianism and Rational Conversion

Introduction

Bayesianism is criticized (most prominently by van Fraassen) for its alleged inability to model a type of rational update which we'll call 'conversions'. Conversions play an important role in the history of science, religion, and thought more generally, and classifying them as outside the bounds of rationality is a big cost for a normative epistemology to bear. For this reason van Fraassen rejects Bayesian Updating, the claim that the only rational way to update is by conditionalizing. But van Fraassen and others who impugn Bayesianism for failing to model rational conversions are mistaken: in the form I have defended it, it can model them easily, and moreover in a way which conveys important insights about their nature. There is thus no need to reject Bayesian Updating on this account. However, van Fraassen's criticism calls much needed attention to a claim with which Bayesian Updating is often, and often implicitly, aligned, which does prohibit conversions from being rational. The claim in question pertains to the *assigning of credence to an agent's evidence proposition*, a step in the updating process which I've called the *input step*. The input step deserves as much philosophical attention as its more glamorous sibling, conditionalization. Confusion about Bayesianism's ability to model rational conversions is just one of many issues on which attention to the input step can shed light.

Rational Conversions

There is a type of update which the literature calls 'radical conversions', 'scientific revolutions', 'paradigm shifts', or 'Gestalt switches'[153]. I will use the term 'conversion' for such an update.[154] Van Fraassen argues that Bayesian Updating should be rejected because, he alleges, it cannot model updates of this type.

Either conversions can be rational, or they cannot. Any epistemology which classified them as irrational would thereby assert that many of the most groundbreaking advances in science ought, from an epistemic point of view, never to have happened, that they were no more than blips of irrationality which were luckily vindicated by the truth utility[155] of the posterior credences in which they issued. It would also assert that many religious conversions, or conversions away from religion, were irrational. That conversions can be rational will not be questioned here.[156] By this I do not mean that all historical conversions *were* rational. All I mean is that they can be rational: a theory of rational updating which cannot model them (or cannot model them accurately) is false.

Before determining whether Bayesianism can model rational conversions, we need adequacy criteria for such a model. Although I disagree with van Fraassen's own model (see Chapter 4; they are non-conditionalizing updates), we can employ the pre-theoretic adequacy conditions which he adduces for one.[157] Conversions, he says (2002, 65), are updates which are

> so radical that they are characterized by a remarkable historical asymmetry. From the posterior point of view, the prior can be made intelligible and the change ratified. From the prior position, however, the posterior view was absurd and the transition to it possible but incapable of justification [i.e., by appeal to any of the credences which the agent has before the conversion, KM].

Two characteristic features of conversions emerge from this picture. Van Fraassen calls one the *prospective unintelligibility of the new*: "the new view [i.e., the agent's post-conversion credence distribution, KM] is literally absurd, incoherent, inconsistent, obviously false, or worse – meaningless, unintelligible – within the older view" (2002, 71-71). There is "a chasm between the old and the new" credence distributions (2002, 112). The second characteristic, which van Fraassen calls *the retrospective intelligibility of older views*, is that, in spite of this chasm, agents "must be able to see [their] present as a rationally endorsable continuation of [their] past" (*ibid.*): an agent who has undergone a conversion must be able to look back on it and affirm that it was rational. It is not easy

[153] See e.g. Kuhn (1962), Feyerabend (1975), and van Fraassen (1989, 2002),
[154] I am not distinguishing between individual and collective doxastic attitudes.
[155] See Chapter 1 for a discussion of truth utility.
[156] Though for an alternative view, see Feyerabend (1975).
[157] Chapter 4 did not address whether his model succeeds at meeting the two criteria in question, since I rejected it for its endorsement of strong doxastic voluntarism.

to spell out these criteria without some formal machinery which could prejudice us in favor of a particular model of rational conversion. But given that all of the models to be discussed here are probabilist, I think this worry can be put aside.

Concerning the first criterion, the *prospective unintelligibility of the new*, it is fair to say that a posterior credence distribution D_{NEW} is 'unintelligible' to an agent, before updating to it, if

> (i) D_{NEW} is defined over at least one proposition *p* which is totally new, in the sense that the agent's pre-conversion credence distribution was not defined over *p*; or
>
> (ii) D_{NEW} contains a non-zero credence in a proposition in which the agent has, before the conversion, a credence of 0.

The idea behind (i) is that the post-conversion agent has a credence in at least one proposition which, before the conversion, she could not imagine having a credence in because she did not grasp it. The idea behind (ii) is that the post-conversion agent has a non-zero credence in a proposition *p* which she could not, before the conversion, imagine having, because *p* was completely excluded from her space of doxastic possibilities.

It is important that these two ways of being prospectively unintelligible, (i) and (ii), refer to the unintelligibility of the new credences from the agent's viewpoint before she acquires evidence. If instead the criteria referred to the time after she receives evidence, conversions would pose deep problems indeed for rational explanation. The reason is, they would amount to an agent's finding a credence distribution unintelligible *as* she changes *to* it, and suddenly finding it intelligible in the instant she adopts it. Any evidence would be incidental to the conversion, for receiving it would have nothing to do with prompting the agent to find the posterior credence distribution intelligible. This sort of conversion would fit the model of a Kierkegaardian 'leap',[158] and would by most counts be epistemically irrational. Or at least, I shall assume that it is, since I am confident that rational conversions can be modelled in a less dramatic way.

As for the second criterion, the *retrospective intelligibility of the old* credence distribution, it may be read in either a strong way or a weak one. Van Fraassen himself opts for the strong reading: the post-conversion agent must regard as rational both her conversion *and* her pre-conversion self. The idea is that she takes her radical transition to have been rational, but that she also sees, in light of her present, enlightened state, how her pre-conversion credences could have been rational on the evidence that she had before converting. She might look back on the successes she enjoyed before converting, and might think that they are what she would have expected on the evidence of her post-conversion credence distribution: the latter "allows for an explanation (in its own terms, of

[158] Kierkegaard (1846, 384).

course) of the old [credence distribution's] successes and of how it was possible for the old, faulty, or false [credence distribution] to have succeeded in those ways" (van Fraassen, 2002, 71-72). The old credence distribution enabled the agent to negotiate the world to all appearances successfully, until whatever precipitated the conversion.

But the strong reading of retrospective unintelligibility threatens to narrow the playing field too much: there may well be pre-conversion credence distributions whose past apparent successes are an utter mystery from the point of view of the agent's post-conversion credences; the pre-and post-conversion credences may be incommensurable. A weak reading of the retrospective intelligibility criterion, by contrast, allows conversions between two incommensurable credence distributions to be rational. It says only that the agent must look back and regard the *transition* – the conversion itself – as rational, if not her pre-conversion credences. The agent might look back and think that she was completely unreasonable for having the credences she did, and regard her fortunate conversion to her present state as an injection of much needed reasonableness.

The conviction that pre- and post-conversion credence distributions are sometimes incommensurable is, after all, one of the major reasons why the likes of Feyerabend, and possibly Kuhn,[159] thought that conversions could not be rational, strictly speaking: how does an agent move to a way of representing the world which is in some sense incommensurable with the one she currently has? Historical thinkers were quite confident of propositions we now take to be impossible. And they thought some propositions impossible which we have now come to realize are not only possible but may be true. It is thus hard to see how one could retrospectively put oneself in the mindset of thinking otherwise. One can still, surely, rightly regard one's conversion as rational, even if one no longer takes one's pre-conversion credences to have been rational. Let us, then, allow van Fraassen's retrospective intelligibility criterion to be read either in the strong way or the weak way: the agent can but need not be able retrospectively to endorse her pre-conversion credences, and either way she must endorse the transition as rational.

Let's regard prospective unintelligibility and retrospective intelligibility as criteria for the adequacy of a model of conversions. Note that this may include non-rational conversions, so that fulfilling these criteria is necessary but not sufficient for a *rational* conversion. Since a theory of rational updating should capture all and only rational updates *simpliciter*, any such theory will add that a sufficient criterion is given by these criteria *and* by the specific provisions of that theory. Of course, as a sufficient criterion this will not satisfy anyone who does not antecedently endorse the theory.

[159] See Feyerabend (1975) and Kuhn (1962). For a discussion about whether Kuhn is correctly interpreted as taking conversions to be irrational, see Brown (1983).

Can Bayesianism model rational conversions? Van Fraassen argues that it cannot, but we'll see that it can after all. The Bayesian account of rational conversion meets the prospective unintelligibility criterion, and both the strong and the weak readings of the retrospective intelligibility criterion.

A Hidden Assumption about the Input Step

Here is van Fraassen's argument for denying that Bayesianism can model rational conversions:

1. A model of rational conversions must meet the prospective unintelligibility criterion and the retrospective intelligibility criterion.

2. An account that employs Bayesian Updating cannot meet the prospective unintelligibility criterion.

3. Hence no account which subscribes to Bayesian Updating can model rational conversions.

We have discussed premise 1. Why does van Fraassen endorse premise 2? He complains that Bayesian Updating makes it impossible for an agent to glean any insight from her experiences which is really *new*, in the sense of making it rationally permissible to discard (at least part of) her credence distribution and to replace it with an entirely new one. Bayesian Updating, he says, "requir[es] … new opinion to be logically forced by new experience and prior opinion together" (1989, 174). The result is that the prospective unintelligibility criterion for conversion is not met: The one permitted way to update (conditionalization) does not leave room for any epistemic surprises, because the agent, before the evidence comes in, can already imagine what it will be like and how she will respond, namely, by conditionalizing on it.[160]

But premise 2 is false, at least as it stands. I'll shortly offer as a counterexample a Bayesian account of rational conversion which meets the prospective unintelligibility criterion. Here is why premise 2 is false: Bayesianism comprises just the two claims Bayesian Updating and Bayesian Statics, and these do not entail that agents can never update in ways which would have struck them as unintelligible before the update. To yield this result, additional assumptions are needed. Recall that Bayesianism as it stands is incomplete, requiring supplementation by claims about evidence, propositional grasping, conditionalization, and so forth. If Bayesian Updating prohibits rational conversion, it is only in conjunction with one or more such supplementary assumptions.

[160] Recall the claim, Grasping, defended in Chapter 3, that having a credence in a proposition *p* requires grasping *p*.

The assumption which causes problems for rational conversion is one with which Bayesian Updating is often implicitly coupled, as we'll soon see. Hence, although van Fraassen is wrong in criticizing Bayesian Updating for failing to model rational conversions, his criticism calls attention to an important problem with a common form of Bayesianism. The assumption is problematic because of what it entails about the *input step*, a topic on which Bayesian theories of rational updating say little but which is no less integral a component of updating than adjusting in light of one's new evidence (see Chapter 3). The input step is the first logical step in the updating process (adjusting by conditionalization is the second); in it the agent assigns a credence to the evidence she has acquired. Let's see what this assumption about input is and how it hobbles Bayesian attempts to model rational conversion. After laying it bare, I'll present my Expansive Bayesian model of conversion, which succeeds because it rejects the assumption. The moral, adduced at the end, will be twofold: first, Bayesian Updating provides an adequate model of conversions;[161] second, it is only at their peril that Bayesian accounts focus on conditionalizing, the second step in updating on evidence, at the expense of the first step, assigning a credence to that evidence to begin with.

Restrictive Bayesianism

Chapter 2 introduced a rough and ready definition of 'conditionalizing update':

> A *conditionalizing update* is one in which the agent sets her new credences equal to the value of her old credences conditional on her new evidence. (*Equal*)

Equal, however, can be read in two ways, and one of these ways, which is a signature view of what I'll call *Restrictive Bayesianism*, entails the problematic assumption about the input step which makes it impossible to model rational conversions.

'New' refers the credences which the agent has as the result of an update. But Equal as stated is ambiguous between two readings of 'old'. It may be understood as either

> A *conditionalizing update* is one in which the agent sets her new credences equal to the value of the credences, conditional on her new evidence, which she has before the input step. (*Pre-Input*)

or

[161] This is not 'empirical' adequacy, since we are dealing with ideal rationality, and empirical adequacy pertains to actual cases. It is adequacy for a model of ideal rationality.

A *conditionalizing update* is one in which the agent sets her new credences equal to the value of her credences, conditional on her new evidence, which she has after the input step. (*Pre-Conditionalize*)

To wrap our heads around these claims, recall the updating diagram from Chapter 5, modified to facilitate drawing the present distinction.

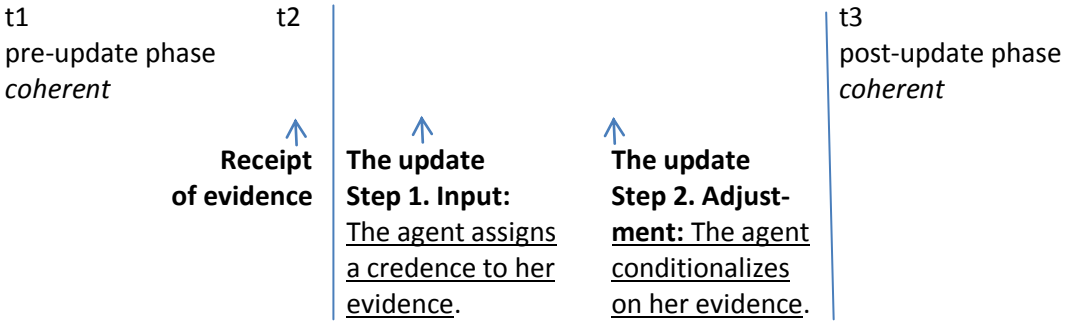

Figure 2: Pre- and post-update, and the two steps of the update.

Pre-Input says that the conditional credences relevant to the agent's conditionalization are the ones he has before Step 1 of the update.[162] Pre-Conditionalize says that the credences in question are the ones he has before Step 2. The 'pre' and 'post' here refer to logical priority, not temporal priority, since the updating process may be instantaneous. Restrictive Bayesianism endorses Pre-Input, which, I'll shortly show, prevents Bayesianism from modelling rational conversions.

But first, the contrast between Pre-Input and Pre-Conditionalize might seem a distinction without a difference: doesn't each just give a different description of the same conditional credences? One might argue as follows: First, any conditional credences which the agent has before she assigns a credence to her evidence are also credences which she has before she conditionalizes, since the input step logically precedes conditionalization. And second, any conditional credences which she has before she conditionalizes are also credences which she has before she inputs a credence to her evidence, since the input step doesn't change an agent's conditional credences – it only changes an agent's (unconditional) credence in her evidence proposition. The distinction between Pre-Input and Pre-Conditionalize, then, seems at best arcane and at worst confusing.

The first point is true, but the second is not. It is not the case that any credences which the agent has before conditionalizing are also credences she has before the input step. There is a possible counterexample, given by a situation in which either the evidential experience or the input step include *bringing about* conditional credences where there were none before. As Chapter 5 has

[162] Both Pre-Input and Pre-Conditionalize may have to add a condition stipulating the earliest credences on which the agent may conditionalize, a feature I'll ignore here.

argued, evidential experiences can bring agents to acquire new conditional credences, and there is nothing irrational about it. I'll develop this point more slowly in the following; this is just to advertise that Pre-Input and Pre-Conditionalize do not necessarily refer to the same conditional credences: not all pre-conditionalization credences are also pre-input credences.

I'll go on to argue that a Bayesian theory of updating which incorporates Pre-Input cannot model rational conversion, but a theory which incorporates Pre-Conditionalize can. This distinction opens the way to showing that van Fraassen's criticism does not succeed against Bayesianism *simpliciter*, but only against a form of Bayesianism which assumes Pre-Input. Van Fraassen's criticism is in this limited sense well taken, though not devastating to Bayesianism. I'll first show how assuming Pre-Input precludes a model of rational conversion; then I'll show that espousing Pre-Conditionalize enables a Bayesian theory to model conversions nicely indeed, and as a bonus, in a way that casts much needed light on their nature.

Pre-Input vs. Pre-Conditionalize: Returning to Bayes's Theorem

Whether one espouses Pre-Input or Pre-Conditionalize depends on how one thinks that Bayes's Theorem should be applied. Recall that this formula governs the way in which conditionalization proceeds. But it is only a formula, and requires supplementary instructions for guiding an agent in using it to update once she has acquired evidence. Pre-Input and Pre-Conditionalize give instructions which are not always compatible. To see this, recall the theorem:

Bayes's Theorem

$$\mathrm{Cr}(P|E) = \frac{\mathrm{Cr}(E)\mathrm{Cr}(E|P)}{\mathrm{Cr}(E)}, \qquad \text{where } \mathrm{Cr}(E) > 0$$

Cr(H) is the agent's credence in a proposition H; Cr(E) is the agent's credence in her evidence; Cr(H|E) is the agent's credence in H conditional on her acquiring evidence E.[163] The caveat ('where Cr(E) > 0') makes explicit that, because it is impossible to divide by zero, there is no defined value for Cr(H|E) when Cr(E) is not greater than zero. Hence for purely mathematical reasons, conditionalization is possible only when the caveat is met.

Of our two interpretations of Equal, Pre-Conditionalize is minimalist: as long as the caveat is met at some point logically prior to the update, the agent may use the formula to conditionalize. Pre-Input, on the other hand, and hence Restrictive Bayesianism, *extends* the caveat by stipulating that,

[163] That is, it is the agent's credence in her evidence whenever this credence is 1; when it is less than 1, the formula for Jeffrey Conditionalization applies instead, but whether Jeffrey or straight conditionalization is employed does not affect the present point.

in order for a conditionalizing update to be permissible, it must be fulfilled before the agent even assigns a credence to her evidence. Hence, according to Pre-Input, regardless of whether the agent's post-input but pre-conditionalization credence in E is greater than zero, if it was not *also* greater than zero before she assigned a new credence to E, then no update is permissible.

Whereas Pre-Conditionalize stipulates the bare minimum for enabling the mathematics of the ratio formula to yield a defined value, Pre-Input goes beyond this by building in a philosophical premise about what sort of evidence can rationally be conditionalized on: only evidence in which the agent's credence was greater than zero *before* she assigned it a new credence in the input step. Thus a Bayesian theory of updating which subscribes to Pre-Input[164] entails this claim:

> When the credence which the agent has in her evidence E before the input step is not greater than zero, the agent may not rationally update on E. (*Restrictive*)

Restrictive is a normative constraint on the circumstances under which an agent may update her credences on assigning a new credence to her evidence proposition. Let's call a Bayesian theory which subscribes to Restrictive *Restrictive Bayesianism*.

Restrictive Bayesianism makes it possible for a rational dilemma to arise. if an agent assigns a non-zero credence to evidence in which her pre-input credence was not greater than zero, there will be nothing she can rationally do to update on it. The first horn is this: If she conditionalizes, she violates Restrictive and hence does something irrational. The second is this: If she does nothing, she has failed to update on her evidence, and violates clause (i) of Evidentialism (see Chapter 1), that the receipt of new evidence rationally mandates an update.[165] And of course, doing something other than conditionalizing, such as rejecting one's evidence, is not a rational option, thanks to Bayesian Updating.

The Restrictive Bayesian opts for the first horn, claiming that an agent who conditionalizes is irrational for violating Restrictive. This horn issues in contradiction, since Bayesian Updating says that conditionalization is necessary for rational updating. A better course of action would be to pre-empt the dilemma's ever arising, which he can do by adding the following principle:

[164] That is, Bayesian Updating, Bayesian Statics, and Pre-Input.

[165] One might wonder whether my framework allows the agent to circumvent this dilemma by assigning her evidence proposition the same credence which she assigned it as evidence – on my framework this qualifies as an update (see Chapter 3, Part II). This suggestion is not relevant here, however, since, here, the agent has already assigned a non-zero credence to her evidence proposition.

> An agent may not rationally assign a non-zero credence to an evidential proposition E if her pre-input credence in E was not greater than zero. (*Restrictive Input*)[166]

Restrictive Input places a normative constraint on what credence an agent can assign to evidence which she has newly acquired. Since newly acquired evidence amounts to no more than a proposition delivered up by an experience, Restrictive Input constrains how the agent is entitled to respond to the deliverances of experience.

At first blush this might seem like a terribly odd thing for a theory of updating to do. But restrictions on the input step are not inherently problematic.[167] The problem is the particular norm, Restrictive Input, endorsed by Restrictive Bayesianism. Restrictive Input, as I'll show momentarily, prevents agents from undergoing rational conversions. But before arguing this, let's ask why one would posit this norm at all. The main reason, as we have seen, is that it enables the Restrictive Bayesian to avoid the consequence that there are situations in which agents may not update on evidence received. Another reason, if not to posit it then at least not to mind that it works well with one's theory of updating, is that one endorses a very strong thesis about the theory-ladenness of observation.

The thesis that observation is 'theory-laden' says that the way in which an agent assimilates new observations of reality is shaped by the representation she has before she makes those observations. On a coarse-grained framework, this is the thesis that the beliefs an agent forms on the basis of experience are in part determined by her prior beliefs. On a fine-grained framework, it is the thesis that the credences which an agent invests in her evidence are influenced by the credences which she has before the experience even occurs.[168] In other words, her confidence that such-and-such a representation of reality is accurate is at least in part determined by the degree of confidence which she already has. If for example an agent is highly confident that she is in a forest, an experience of being appeared to by a tall and branchy object will cause her to assign a high credence to the evidence that there is a tree before her. If however she highly confident that she is in a laboratory being injected with hallucinogenic drugs, she will assign a significantly lower credence to this proposition. Surely observation is to some extent theory-laden. If this is so, then surely the credence which an agent assigns to his evidence is influenced by the credences which he has beforehand.

---

[166] Of course, the agent can acquire such evidence if she doesn't need to change any other credences in light of it, but we are interested only in cases in which the evidence does make an agent's credence distribution incoherent.

[167] See Chapter 3 for arguments that the input step can and should be constrained in some respects.

[168] For original statements of strong views about the theory-ladenness of observation, see Popper (1963), Lakatos (1978), or a somewhat more subtle view, Sellars (1956). For helpful discussions of theory-ladenness see Brewer and Lambert (2001) and Kosso (1992), Chapter 6.

There are stronger and weaker claims about theory-ladenness. A particularly strong variety is this:

> An agent for whom a proposition *p* is doxastically impossible at a time t cannot invest a non-zero credence in *p* at t or any time thereafter. (*Nomological Impossibility*),

where the modality of 'cannot' is nomological: the laws which obtain in our universe[169] ensure that an agent cannot invest a non-zero credence in *p*.[170] Nomological Impossibility says more than that what one does with one's evidence *e* is *influenced* by the credence which one has in *e* before one receives *e* as evidence: it says that there are cases in which prior credence renders an agent *incapable* of forming particular credences in her evidence, and that such cases *always* arise when her credence in her evidence proposition was, before she assigned it any input credence, non-zero.

If the descriptive claim Nomological Impossibility were true, then its normative cousin, Restrictive Input, would always (in our universe) be trivially satisfied: agents would always form the credences in their newly acquired evidence which they would be rational in forming.[171] Perhaps it is commitment to Nomological Impossibility, under the guise of a strong thesis about the theory-ladenness of observation, which makes Restrictive Bayesians sanguine about accepting Restrictive Input. But Chapter 5 has argued that Nomological Impossibility is false: although there may be *some* cases in which an agent is incapable of moving from a zero to a non-zero credence in some proposition she acquires as evidence, it is surely not a universal feature of such cases. Since, then, Nomological Impossibility is false, Restrictive Input is by no means a trivial norm. If agents (nomologically) *can* invest a non-zero credence in evidence in which their pre-input credence was not greater than zero, then there are possibly cases in which agents become irrational by doing so.

Why endorse a norm that has this result? Restrictive Input is perhaps plausible when one has in mind the pictorial model of updating which van Fraassen calls the 'muddy Venn diagram' (1989, Chapter 13, introduced in Chapter 5). Imagine a Venn diagram of which each cell represents a set of doxastic possibilities. When a cell has mud on it, the possibilities within it are live for the agent, and when the mud is removed, so are the possibilities. On this model, becoming certain of an evidence proposition E is represented by removing any mud from the cell representing not-E, and assigning E an uncertain credence is represented by the shifting of some but not all of the mud from E to not-E

[169] The proponent of Nomological Impossibility need not be committed to realism about laws of nature; such laws might be descriptions of properties of the universe, rather than rules which govern in some more robust sense.

[170] This is more than the claim that *p* will not actually obtain; it is the claim that, because of the properties of our universe, *p cannot* obtain in our universe. Features of the universe which might make Nomological Impossibility true include such things as the cognitive properties of human agents or the properties of credences.

[171] This is assuming that there are no other restrictions on the evidence an agent may rationally acquire.

or vice versa. Conditionalization is represented by the redistribution of the removed mud over the remaining cells, in proportion to the amount of mud which those cells already had. On the muddy Venn model evidence can do only one of two things, depending on whether the agent assigns it a credence of 1 or of less than 1. If she assigns it a credence of 1, then her space of doxastic possibilities *shrinks*: possibilities inconsistent with her evidence are deleted from the diagram when all of the mud is removed from them; there can never be mud on them again. (The removed mud is then redistributed over the remaining cells in proportion to the amount of mud they had already.) If the agent forms a credence of less than 1 in her evidence, then the mud is redistributed among all of the cells which still have mud on them.

Thus on the muddy Venn model of updating, the agent's posterior credence distribution is always a sub-model of the credence distribution which she has before assigning any credence to her evidence. If this model is taken to be normative over updating, then there is no room for the agent either to 'recover' deleted cells or to 'create' cells where there were none before, for both would violate Pre-Input. Restrictive Input seems a natural way of fleshing out this initially plausible picture.

But of course, a picture in this context is only a good heuristic if it is true to reality in the relevant ways. If in fact agents *can* invest non-zero credences in evidence in which their credence was not previously greater than zero, and if they can do so rationally, then the muddy Venn picture of ideal credence formation misses its mark. It becomes a hindrance to understanding, and the picture must be discarded or adjusted. That Restrictive Bayesianism – and thus the muddy Venn picture – cannot model rational conversion provides strong reason for doing one or the other, as we'll now see.

Restrictive Bayesianism and Rational Conversion

Restrictive Bayesianism can model conversions, but it cannot pronounce them rational. It is the first criterion, the prospective unintelligibility of the new credence distribution, on which Restrictive Bayesianism stumbles.

Recall that there are two ways of fulfilling this criterion. The first is for the new (i.e., post-update) credence distribution (call it 'D_{NEW}') to be defined over at least one proposition *p* which is totally new to the agent, in the sense that the credence distribution which the agent had before the input step (call it 'D_{OLD}') was not defined over *p* at all. The second way of fulfilling the prospective-unintelligibility criterion is for D_{NEW} to contain a credence which is outside the range of credences which were, before the evidential experience, doxastically possible for the agent; i.e., to contain a non-zero credence in a proposition in which her D_{OLD} credence was 0. In both cases Restrictive Bayesianism says that there is no rationally possible update to D_{NEW}: Bayesian Updating forbids non-

conditionalizing updates, and Restrictive forbids an agent from conditionalizing on a proposition in which she does not, before the input step, have a credence which is greater than zero (criterion (i) for prospective unintelligibility); this includes propositions in which she had no defined credence at all at that stage (criterion (ii)). Hence Restrictive Bayesianism cannot pronounce any conversions to be rational at all.

Let's return to van Fraassen's argument from rational conversion. I've just shown that, if we substitute 'Restrictive Bayesianism' for 'Bayesianism', this argument is sound. But how does it fare as it stands, which is against *any* Bayesian theory of updating? As we'll now see, premise 2 is shown false by the counterexample of a Bayesian theory which denies Pre-Input and instead endorses Pre-Conditionalize.

Expansive Bayesianism

Let's return to the Muddy Venn model, which we now see must either be discarded or adjusted. I'll adjust it in a simple way: by permitting 'deleted' cells to be recovered or completely new cells to materialize. But for this to happen they need some mud, which must be re-allocated from the rest of the diagram. If the new cell is a previously deleted but now recovered cell (call it 'P'), then its mud comes, straightforwardly, from the cell named 'not-P', to which a coherent agent will allot 100% of his mud. If the new cell is making its debut on the diagram, representing a proposition which the agent did not grasp before she acquired it as evidence, then it will come into being simultaneously with a cell representing its negation. Where is the mud to come from to cover one or both of these cells? It comes from the cells which are already mud-covered which represent propositions incompatible with either of the propositions represented by the new cells. Let's call this the *expanded muddy Venn diagram*.

The expanded muddy Venn diagram models a Bayesianism which subscribes to Pre-Conditionalize, which I call 'Expansive Bayesianism'. Expansive Bayesianism does not entail Restrictive and hence has no need for any such stipulation as Restrictive Input, which it denies. The denial of Restrictive Input says that:

> An agent may rationally form a non-zero credence in an evidential proposition *e* if her credence in *e*, before receiving it as evidence, was not greater than zero. (*Not-Restrictive Input*)[172],

[172] Of course, the agent can acquire such evidence if she doesn't need to change any other credences in light of it, but we are interested only in cases in which the evidence does make an agent's credence distribution incoherent.

Not-Restrictive Input entails, of course, that it can be rational for an agent to assign (i) a non-zero credence to a proposition which, before the input step, was doxastically impossible for her, or (ii) in a proposition which, before she acquired it as evidence, she did not grasp.

Expansive Bayesianism is compatible with a much more natural interpretation of Bayes's Theorem and its caveat. It constrains the agent no more than is necessary to satisfy the mathematics of the theorem. Thus it requires only that the agent's *pre-conditionalization* credence in her evidence proposition be greater than zero. The credences which an agent has before assigning a credence to her evidence do not determine the credence which she assigns. The agent is free to let experience deliver up surprises. Such surprises are what trigger rational conversions, according to Expansive Bayesianism.

The sort of epistemic surprise which can bring about rational conversion comes in two varieties, each of which corresponds to one of the criteria for prospective unintelligibility.

Impossible-Evidence Conversions

Corresponding to criterion (ii) for prospective unintelligibility is an experience which re-introduces a proposition which, before the experience, was doxastically impossible. On the muddy Venn diagram, recall that this is impermissible: once a cell has had all of its mud removed, it is deleted from the diagram and cannot be re-instated. On the expanded muddy Venn diagram, however, which is the alternative picture I am endorsing, there is no such prohibition. A cell is not deleted when its mud is removed; it remains, clear of mud. Thus an evidential experience, which – however surprisingly surprisingly – presents an agent with the proposition corresponding to that cell can induce the agent to restore some mud to it. When this happens, we'll call the surprising evidential proposition *impossible evidence*, shorthand for 'evidence which was doxastically impossible before the evidential experience'. An *impossible-evidence experience* is one in which the agent acquires impossible evidence. Such an experience can issue in an *impossible-evidence conversion*. In an impossible-evidence conversion, an agent updates on a piece of evidence which she never imagined she could receive. She can do so by assigning some non-zero credence to it, and conditionalizing on it. But in order to conditionalize on her impossible evidence, she needs credences conditional on it. Where do these come from?

There are two ways, discussed in Chapter 5, in which an agent can acquire credences conditional on her impossible evidence.

(1) First, she may already have them. Recall that, on my view, conditional credences are basic, not defined in terms of unconditional credences. There can thus be well defined credences given some

proposition *p* even if an agent's unconditional credence in *p* is zero. Thus, even though a proposition *p* is doxastically impossible, the agent may still have credences about what things would be like were *p* actual. If the agent is certain that the moon is not made of green cheese, she may nonetheless have credences about how likely it would be that other celestial bodies are made of green cheese, given that the moon is. She might reason by a principle of statistical inference to the effect that, if one of a certain kind of thing is *F*, then there is a greater probability that other things of that kind are also *F*. Hence if she already has credences conditional on her impossible evidence, these will come into force when that evidence becomes possible for her.

(2) But the agent need not have such conditional credences already. I submit (and have argued in Chapter 5) that evidential experiences can help bring them about, and that they may also be the products of agents' deliberation.

Impossible-evidence conversions are the most dramatic when the agent's credence in *p* moves from zero to very high. But the update is no less of a conversion even when the agent's new credence in *p* is very low: *any* non-zero credence in *p* is doxastically impossible for the agent from her pre-experience viewpoint.

Impossible-evidence conversions meet both criteria for conversion, prospective unintelligibility and retrospective intelligibility. We've just seen how they account for the prospective unintelligibility of the agent's post-update credences: they are updates on evidence which she never imagined it possible that she could receive. As for the retrospective intelligibility of the pre-experience credences, impossible-evidence conversions meet this criterion on both its strong and its weak readings. Recall the strong reading, according to which the post-conversion agent looks back on her pre-conversion credences *and* the conversion from them as rational. Expansive Bayesianism can account for this: The post-conversion agent can say something like the following of her pre-conversion self: "It is no wonder I had a zero credence in *p*; I got along just fine without it until the experience as of *p*. Certainly, this or that aspect of reality remained unexplained for me, but how could I have suspected that my zero credence in *p* was to blame, especially," she might add, "when I chalked up any sense of confusion my own dimness?"

On the second reading of the retrospective intelligibility criterion, which requires merely that the agent regard the *transition* from her prior credences as rational, even if she does not regard those credences themselves as rational, Expansive Bayesianism equally has no trouble accounting for retrospective intelligibility: The agent can say: "My past credences were thoroughly unreasonable; I can't imagine anymore what it was like to have them, since I can't imagine what it was like to think that such-and-such was completely and utterly impossible. Fortunately I had an evidential experience which showed me the light and enabled me to update out of my benighted state." Hence

impossible-evidence conversions meet both criteria for rational conversion. Expansive Bayesianism is a successful counterexample to van Fraassen's premise 2.

Undefined-Evidence Conversions

There is a second form of rational conversion which Expansive Bayesianism models as well and which corresponds to criterion (i) for prospective intelligibility. This form of conversion, which we'll call *undefined-evidence conversions*, is brought about by a second variety of epistemic surprise which an evidential experience can deliver up: the coming to grasp of a new proposition. We'll call such a proposition *undefined evidence*, shorthand for 'evidence in which the agent did not have a defined credence before the evidential experience'.[173] The agent may or may not invest a credence in the newly grasped proposition – she has the rational option of suspending judgment toward it. But if she does invest a credence in it, then the set of propositions over which her credences are defined expands, or, more accurately, becomes a different set, which is larger than the one she had before. This is what Chapter 1 called a 'C-set expansion'. The muddy Venn diagram makes no provision for the receipt of undefined evidence: just as it does not allow a mudless cell to recover any mud, neither does it allow any new cells to be generated. But the expanded muddy Venn diagram makes no such prohibition. The expanded set is the set of propositions over which her pre-experience credences were defined plus the new proposition and its truth-functional combinations.

Undefined-evidence experiences put an agent in a position to update on a piece of evidence which she never imagined she could receive because she could not grasp it. She need only assign it some non-zero credence and conditionalize on it. But again – where do her conditional credences come from? In the case of undefined-evidence conversions, there were two possibilities. One was that the agent could have conditional credences 'waiting in the wings'. But in undefined-evidence cases, this is not an option, because the agent has only just come to grasp the evidential proposition, and hence (because having a credence in a proposition entails grasping that proposition) could not have had credences conditional on it before coming to grasp it. The second possible source of new conditional credences, in impossible-evidence conversions, is the evidential experience itself. It is plausible to suppose that an experience which generates evidence can also generate credences conditional on that evidence, in the event that the evidence is in some way surprising. This option is available for undefined-evidence conversions. It stands to reason that, in coming to grasp a proposition for the first time, an agent will also form credences about how likely that that

[173] Another kind of undefined-evidence scenario which I do not discuss here is one in which the agent did grasp her evidential proposition before having the evidential experience, but simply suspended judgment from it until then.

proposition makes other propositions. She thus converts by updating, on her ungrasped evidence, in accord with the new conditional credences.

One might think that undefined-evidence conversions are plausible to a point. But why, one might wonder, should an agent who has a zero credence in a proposition *p* before having an undefined-evidence experience suddenly change her credence in *p* to non-zero, simply because she has come to grasp a new proposition? The answer is this: she will not change her credence in *p* simply because she has come to grasp the new proposition. She will do so in the event that the evidential experience *also* generates credences in *p* conditional on her newly grasped evidence which are greater than zero. She may then conditionalize in accord with these; if she does not acquire new conditional credences in this way, then no conversion need result.

Undefined-evidence conversions meet both the prospective unintelligibility criterion and the retrospective intelligibility criterion. It is clear how the former is met: the agent cannot grasp her evidence before she acquires it – of course she cannot imagine ever reasoning to any conclusion from it. But after she acquires the proposition as evidence, she can grasp it and hence more possible credence distributions become doxastically open to her than she could, before the evidential experience, have imagined.

As for the retrospective intelligibility criterion, undefined-evidence conversions can fulfill both its strong and its weak readings. On the strong reading, which requires that the post-conversion agent regard her pre-conversion credences as rational, the post-conversion agent can say of her pre-experience self: "It is no wonder I had a low credence in *p*; I hadn't even yet grasped *e* let alone had a high credence in it. Now, having my credences defined over this completely new proposition *e* (as well as its truth-functional combinations), I have more information than I did then; given the benighted state of my pre-experience self, it makes perfect sense that I had such a low credence in *p*". On the weak reading of this criterion, according to which the agent need merely regard the *transition* as rational, she can say: "My past credences were thoroughly unreasonable; fortunately, coming to grasp *e* gave me a rational transition out of that benighted state".

Hence Expansive Bayesianism presents an adequate model of rational conversions. Whether they are impossible- or undefined-evidence conversions, they fulfill both pre-theoretic desiderata for conversion.

Light Shed on the Nature of Rational Conversion

Expansive Bayesianism is more than adequate. It offers insight into the nature of conversion itself. This is an achievement: the puzzle of how conversions can be rational is subject to much

philosophical scrutiny, which has driven the likes of Kuhn, Feyerabend, and Kierkegaard to hold that they are neither rational, nor even a-rational, but instead positively *ir*rational. This is a large bullet to bite, as we have seen, since it implies that many of the most groundbreaking realizations in science, religion, political theory, and so forth, have been irrational. Expansive Bayesianism shows that conversions can be rational after all: they can proceed by conditionalization. This is important for two reasons. First, it satisfies evidentialist inclinations: rational conversions are rational because they are appropriate responses to evidence. Second, it accounts for the radical nature of conversions: The evidence which precipitates a rational conversion is evidence which the agent could never, before the evidential experience, have imagined receiving.[174]

Thus van Fraassen's argument against Bayesianism fails, since Expansive Bayesianism falsifies its second premise. But his argument nonetheless does the great service of pointing out the conversion problem, which we have discovered arises from a common but highly problematic assumption about conditionalization. This assumption motivates what we have seen to be a wrong-headed but usually implicit constraint on *the credences which one can rationally assign to one's evidence*. The constraint is Restrictive Input, given above, which forbids an agent from assigning a non-zero credence to an evidential proposition which, before she acquired the evidence, was either zero or undefined.

Hence there is a second moral we can draw from our discussion of van Fraassen's attack on Bayesianism: a theory of rational updating must not focus so closely on the second step, conditionalization (more generally, adjustment), that it loses view of the first step, the assigning of a credence to one's evidence.

The Objection from 'Seemings' Circumvented

Recall Chapter 3's argument against Seemings, the claim that a representational experience presents a rational agent with propositions about how things seem, and not with propositions about how things are. The present chapter reveals why it is so important for my project to establish the falsehood of Seemings.

The reason is that the Expansive Bayesian view of rational conversion will be limited if an agent's evidence consists only in propositions about how things seem to her, rather than about how they are independently of her experiences. If Seemings is true, then agents will only ever receive as evidence propositions about how things seem to them; the only way to change their credences in propositions about how things are, on this view, would be by inductive inference from propositions

[174] Of course, agents *can* imagine receiving as evidence a general proposition such as *There is a piece of evidence E which I can't at present grasp but which, if I could, would alter my credences vis-à-vis the proposition P*, which is of course entailed by E.

about how things appear. This would limit the Expansive Bayesian account of rational conversion in the following way:

If Seemings is true, impossible-evidence scenarios will be much rarer, although they would still exist. They would be rare because, if evidence consisted only in propositions about seemings, an agent would be barred from receiving impossible evidence about how things are. Thus an agent with a credence of zero in *A pig is flying* cannot receive this proposition as evidence, but instead only the proposition *It seems to me as if a pig is flying*. The reason is, if she has an experience as of a pig's flying, her zero credence in *A pig is flying* will be plugged into the numerator of Bayes's Theorem, and hence the resulting posterior credence in *A pig is flying* will always be zero, since multiplying and dividing by zero will always result in zero.

But Chapter 3 argued that Seemings is false, so I do not take it to present a viable objection to the Expansive Bayesian account of rational conversion.

Even if such an extremely internalist view of evidence as Seemings were true, however, the Expansive Bayesian account of rational conversion would not be vitiated. The truth of Seemings would not affect scenarios in which the agent receives previously *undefined* evidence: if she comes to have an evidential experience in which she grasps a new proposition to which she then assigns a non-zero credence, she may still undergo a rational conversion. Nor does the worry about Seemings touch all impossible-evidence scenarios. In particular, it does not affect situations in which the agent receives, as evidence, a proposition *about how things seem to her* which, before the experience, was doxastically impossible. If before an evidential experience she had a zero credence that it would ever appear to her as if a square is round, and if she then has such an experience and assigns a non-zero credence to *It seems to me as if a square is round*, she might conditionalize on this seemings-proposition to arrive at a non-zero credence in *There is a round square*. Thus impossible-evidence scenarios might still arise, although presumably much more rarely, as would undefined-evidence scenarios. The truth of Seemings would not vitiate the Expansive Bayesian account of rational conversion, but it would limit it significantly. Fortunately, Seemings is false.

Conclusion

Restrictive Bayesianism's failure to model rational conversion can be traced to its endorsement of Pre-Input, above, which understands conditionalization as making one's posterior credences equal to the value of one's old credences conditional on the evidence received, where 'old' is cashed out as 'before the agent assigns a credence to her evidence'. Pre-Input entails that it is irrational to update on evidence which, before one has received it from a representational experience, was

impossible or undefined. Hence, to ensure that such evidential input cannot arise to begin with, Pre-Input strongly motivates a prohibition on assigning non-zero credence to any such evidence at all. Pre-Input is a piece of philosophical ideology motivated by an overly narrow picture of the updating process.

But we saw that Pre-Input goes beyond the pure mathematics of the ratio formula, which merely needs the agent to invest a non-zero credence in her evidence proposition. Pre-Conditionalize, by contrast, which understands conditionalization as the making of one's posterior credences equal to the value of the credences which one has *after* one assigns an input credence, sets experience free to do its proper job in the updating process. Experience can surprise us by providing us with evidence which we could never, before our evidential experience, have imagined receiving.

7. Bayesianism and Rational Permissiveness

Introduction

Some liberationists reject Bayesian Updating, the claim that the only rational way for an agent to update her credences is by conditionalization, on the grounds that it is incompatible with a view, which I'll call *Permissive Updating*, which says that rationality sometimes permits the assignment of more than one posterior credence to a proposition on receipt of a given piece of evidence. Permissive Updating belongs to what I have called a 'liberationist' picture of rationality, which emphasizes rational permission rather than obligation. I'll argue first that Permissive Updating is harder to establish than meets the eye, and second, that even so Bayesian Updating, perhaps surprisingly, is consistent with it: there are scenarios in which Bayesian rationality may permit the assignment of more than one posterior credence on receipt of some (total) evidence. To take up the theme these pages are developing, Bayesianism is a much more flexible theory of rational updating than many think.

Permissiveness About Updating

Rational updating is either permissive or directive. It is *permissive* just in case, once an agent assigns an input value to her evidence, there are two or more posterior credences in some proposition *p*

which she may rationally adopt in response to that evidence.[175] Rational updating is *directive* just in case it is not permissive, that is, just in case there is only one rational posterior credence which an agent may assign to *p* given her evidence. We'll focus on the claim

At least some rational updating is permissive. (*Permissive Updating*)

To zero in on the notion of permissiveness about updating, imagine that an agent has an evidential experience and performs the logically first step of an update, assigning a credence to her evidence proposition.[176] Permissive Updating is the claim that, once she has done this, there are two or more posterior credences which *that agent*, given her unique credences and epistemic history, can rationally adopt on the basis of this input value. Of course only one of these values, at most, would result from the agent's conditionalizing, since we are assuming that she is (weakly) coherent and that credences are precise. Thus some have taken Permissive Updating to provide fodder for the following argument against Bayesian Updating:

Argument from Permissive Updating[177]

1. Permissive Updating is true. (assumption)

2. Bayesian Updating entails not-Permissive Updating. (assumption)

3. Hence Bayesian Updating is false. (from 1, 2)

Let's first look at the first premise. Who cares about Permissive Updating? Liberationsts (such as van Fraassen; see Chapter 4) care about it a great deal. One of van Fraassen's concerns stems from his voluntarism about updating, the claim that sometimes agents have the psychological ability to choose between two or more posterior credences on receipt of some piece of evidence; Chapter 4 argued that this strong form of doxastic voluntarism is false. But Permissive Updating does not entail that updating is voluntary: options need not be subject to voluntary choice in order for them to be rationally permitted. There are other reasons, at least prima facie ones, for endorsing Permissive

[175] There is another way in which rational updating may be permissive, that is, if there are two or more rational credences which an agent may assign to her evidence itself. But this form of permissiveness is not typically used as an objection to Bayesian Updating. I'll return to it at the end of the chapter, but for now we will focus on permissive updating as I have characterized it.

[176] Recall that conditionalizing updates have two logically (if not temporally) distinct steps. The first is the input step, in which the agent assigns a credence to her evidence proposition *e*. The second is the conditionalization step, in which she adjusts her other credences so that they cohere probabilistically with her new credence in her evidence.

[177] The main proponent of an argument of this form, targeted in particular against Bayesian Updating, is van Fraassen (1989), whose version will be discussed later.

Updating. One is that it is entailed by a claim about the nature of evidence; another is that it is required in order to account for epistemic virtues; and a third is that some updating scenarios are too complex to admit of only one rational option. However, I'll show that none of these arguments supports Permissive Updating. Thus Permissive Updating is a weak premise on which to stake the fate of Bayesian Updating.

But there is another reason why Permissive Updating should not be used to argue against Bayesian Updating: it is perfectly compatible with Bayesian Updating. This may come as a surprise; after all, doesn't Bayesian Updating require that agents update by conditionalizing, and doesn't conditionalization yield only one posterior, namely, the posterior borne by plugging the input and prior credence values into Bayes's Theorem? Yes, but as previous chapters have indicated, Expansive Bayesianism adds nuance to this story: there may be situations in which agents do not, on receipt of evidence, already have defined prior credences. The rational posterior in such situations may, for all Bayesian Updating has to say, be up for grabs.

Let's clarify Permissive Updating. First, note that, even if updating is permissive, rationality still requires the agent to adjust her credences in some way on receipt of evidence, even if this is only to change her credence in her evidence proposition and nothing else. The reason is given by clause (i) of Evidentialism (see Chapter 1), that the receipt of new evidence rationally mandates an update. Second, Permissive Updating pertains only to *ideally* rational updating. It does not pertain to the way in which an agent may or may not update who has already fallen short of the ideal. *Regulative* rationality, which governs non-ideal agents, may be very permissive for all I am arguing.

Third, we must carefully distinguish Permissive Updating, the claim that some *updating* is permissive, from the closely related claim that some *evidence* is permissive. Permissiveness about updating is localized to an epistemic agent, whereas a permissiveness about evidence is not. A body of total *evidence* is permissive just in case, when abstracted from any localized agent whose evidence it may be, it rationally admits of multiple interpretations, where an interpretation of evidence is an assignment of a probability value to some other proposition given that evidence. Evidence itself may be permissive in this sense without its being the case that the agent, given her unique arrangement of credences and her unique epistemic history, has permission to update to any of the values which her evidence, in the abstract, may allow for. In a moment we will see how permissiveness about updating and permissiveness about evidence relate logically; for now it suffices to distinguish them.

Arguments for Permissive Updating

Let's look at three arguments for Permissive Updating.

Argument for Permissive Updating from Permissive Evidence

One reason for endorsing Permissive Updating has been offered by White (2005), who argues to it from the claim that:

> At least some evidence is permissive. (*Permissive Evidence*),[178]

where evidence is permissive, recall from the previous section, just in case, in abstraction from any particular agent whose evidence it may be, it admits of multiple rational interpretations. Recall also that an *interpretation* of a piece of evidence is an assignment of a probability value to some proposition or propositions given that evidence. Thus Permissive Evidence differs from Permissive Updating in that the former pertains to which probability values are permitted given a piece of evidence, and the latter pertains to what a localized agent with an epistemic history and unique credence distribution is permitted to do with her credences.

Many Bayesians endorse Permissive Evidence. The most extreme hold that there are no rational constraints on credences (including conditional credences) other than coherence, which means that there are multiple different rational credences for a proposition P conditional on some body of evidence E. But even among Bayesians who reject Orthodox Statics and hold there to be some rational constraints on conditional credences, few Bayesians of whom I am aware, beyond the self-styled 'Objective Bayesians',[179] take there to be only *one* rationally permitted conditional credence for each proposition/body of evidence pair. Thus Permissive Evidence is germane to Bayesianism, and endorsed by many Bayesians, though not required by it. More than this, Permissive Evidence is plausible. It seems crazy to stipulate that a total body of evidence would legitimate a credence in some proposition P of, say, .23 but not one of .24. There may be some restrictions on what interpretations of evidence are permitted, but surely, for each proposition-evidence pair, there is at least a small range of rational options.

[178] Permissive Evidence is a compact equivalent of the negation of what the literature has come to call 'The Uniqueness Thesis', coined by Feldman (2007) and formulated by Kelly (2010) as: "For a given body of evidence and a given proposition, there is some one level of confidence that it is uniquely rational to have in that proposition given that evidence". The Uniqueness Thesis is a claim about rational credence at a time: about what credence is rational to have on the assumption that one also has such-and-such a body of evidence (or which it would be rational to have were one to have that evidence).

[179] See e.g. J. Williamson (2009).

But White argues that, if Permissive Evidence is true, it is hard to see how Permissive Updating would not also be true. He presents readers with a series of thought experiments of the following form: Suppose that evidence E (which receives an input value of *z*) permits a probability assignment of *x* to P as well as an assignment of *y* to P. White suggests that an agent who receives E (and assigns it a credence of *z*), and who updates to a credence of *x* in P, can, without additional evidence beyond E, rationally change her credence in P to *y*. That is, she can rationally 'slide', without additional evidence, from one credence permitted by her evidence to another. For opponents of Permissive Updating, sliding updates are permissive in the worst way: not only does the agent move between two rational posteriors, she does so having *already* opted for one and changed her mind. In Bayesian terms, this amounts to a seriously non-ideal move, since it is a non-conditionalizing update: a slide amounts to the agent's switching her conditional credences without being prompted by evidence to do so. But if Permissive Evidence is true, White argues, it is hard to see how permissive updates could not be rational, even when such updates are slides. He takes this consequence to be bizarre and counterintuitive. And Bayesians should, too, since slides are non-conditionalizing updates. He suggests in conclusion that he can see little way to avoid Permissive Updating, short of abandoning the eminently reasonable Permissive Evidence.

What exactly is White's claim? It cannot be that Permissive Evidence *entails* Permissive Updating. A counterexample is given by a world in which Permissive Evidence is true and a rule which forbids permissive updating applies. Liberationists think that Bayesian Updating is such a rule, but we'll see that they are wrong. But Bayesian Updating does prohibit sliding permissive updates; hence a world in which Bayesian Updating is true is one in which slides are irrational, even if not all permissive updates are. So what is White worried about?

What White's thought experiments seek to impress on us, I take it, is the intuitive oddity of thinking that there are norms about updating which are more restrictive than facts about evidence. Why should evidence admit of multiple interpretations, yet an agent be prohibited from availing himself of more than one of them? Norms about updating, the thought seems to be, should be determined by facts about evidence. Take an analogy. Raw honey is healthy to eat and, because of its enzyme content, good for topical application in cleansing one's skin. Raw honey is like permissive evidence in that two (let's suppose) equally good things can be done with it. It seems completely ad hoc, then, to mandate that people only ever eat raw honey and never clean their skin with it. What would be the point of such a mandate? If anything it would deprive many of the cleansing effects of a potent natural substance.

The relationship between permissive evidence and updating, on White's view, seems similar. He seems to be saying that agents should be permitted to do, with their evidence, exactly what the

properties of that evidence dispose it to be used for: assigning one probability, or assigning another. Forbidding one probability assignment while permitting another seems completely ad hoc, and moreover, out of step with the nature of the evidence itself. The ad hocness seems only intensified when we note that one agent, because of his prior credences, will be allowed one permitted probability assignment *x* and forbidden *y*, whereas another agent, because of her prior credences, will be forbidden *y* but permitted *x*. Surely norms for updating are every bit as egalitarian as evidence itself.

But this reasoning is too quick. There is an important difference between evidence and the updates which an agent performs on evidence. Recall the distinction between Permissive Evidence and Permissive Updating. Permissive Evidence is a claim about what probability values a body of evidence confers on propositions not contained in it, regardless of whether those probability values or that evidence belong to some agent. Permissive Updating, by contrast, is a claim about what agents may rationally do with *their* subjective probability values in response to *their* total evidence. The thrust of White's argument is that properties of evidence *completely determine* norms governing the way in which specific, localized agents should respond to that evidence. That is, White seems to assume something along the lines of:

> Norms for rational updating are determined by nothing other than the properties of evidence. (*Evidence Determines Updating*)

But why should Evidence Determines Updating be true? Let's return to our honey analogy. That honey is good for eating and cleansing does *not* entail that every agent should both eat it and cleanse his skin with it. One agent may have a bowel condition which the honey's sugar content will aggravate so that he cannot absorb its nutrients; another agent may have sensitive skin which is irritated by honey. Honey is good for eating and cleansing only *dispositionally*. A disposition is only triggered in specific circumstances. The properties of honey do not determine what agents ought to do with it; contextual factors feature as well.

I submit that something similar holds for evidence and epistemic agents. Evidence alone does not determine how agents ought to update on it: Evidence Determines Updating is false, even if the following modified norm is true:

> Norms for rational updating are determined in part by the properties of evidence. (*Evidence Partly Determines Updating*)

What other factors determine norms for updating? Such factors are given by the local epistemic situation of each agent. One factor, says Bayesian Updating, is what the agent's conditional

credences are. If, on the supposition that E is true, the agent takes a proposition P to be *x* probable, it would surely be irrational for her not to update to *x* on becoming certain of E, but to update to some other permitted value, *y*, instead. She would be contravening her own epistemic lights. Another factor which determines norms for updating is an agent's expected truth utility assignment.[180] Even if a total body of evidence E allows a proposition P to receive probability *x* or *y*, if the agent gains a higher expected truth utility by updating to *x* than to *y* on receipt of E, then surely she would be irrational to update to *y*. This is a second way in which she would be violating her lights.

The situation is slightly more complex if the agent becomes certain that evidence E, which by her conditional credences makes a proposition *p* probable to degree *x*, in fact permits an assignment either of *x* or of *y*.[181] Surely then, White might argue, she may permissibly slide to *y* even if she had already updated to *x*; at least, it is hard to see why she should be forbidden from doing so. Indeed, the intuitive pull of White's thought experiments comes from the fact that we, the readers, are told that an agent's evidence is permissive. If we didn't know this, then our intuitions would probably not (to the extent that they do) push us toward thinking that permissive evidence makes permissive updates rational. If White's thought experiments work on us, they do so because *we* have evidence to the effect that there is more than one rational way to interpret the evidence of the characters in his thought experiments. But this line of reasoning shows White's thought experiments to be weaker than they at first appear. They are premised upon the reader's having more evidence than the characters do, in particular:

E* Evidence E admits of multiple rational interpretations.

Conditional on E*, an agent may well pause – and rationally – about whether to update in accord with her credence in P conditional on E or whether to assign some other permitted value to P. But situations like this do not support Evidence Determines Updating. For the issue in such situations is *not* how to interpret E. It is how to interpret the total evidence *E and E**: Given evidence E and given that E rationally permits multiple interpretations to P, what credences should an agent assign to P? Bayesian Updating says that she should conditionalize her credence in P on her credence in *E and E**.

A second point against Evidence Determines Updating is this: That a body of evidence permits two different credences does not entail that both of the permitted credences have the same expected truth utility for the agent; indeed, they cannot. Arguably a rational agent will update to whichever credence has higher expected truth utility. Perhaps finding out that her evidence is

[180] See Chapter 1 for discussion of this notion.
[181] Other problems arise of course if she becomes certain that it permits neither.

permissive will affect her expected truth utility values, but if so then she should not slide, but should rather conditionalize on her new total evidence which includes this fact.

Thus Evidence Determines Updating is unmotivated. Even if evidence is permissive, rational updates are determined by much more. Permissive Evidence does not furnish an argument for Permissive Updating.

For Permissive Updating: Epistemic Virtues

Van Fraassen offers the following argument for Permissive Updating (1989, 174):

> [W]ho says that we must commit ourselves to [conditionalization] beforehand [KM: i.e., before receiving evidence]? David Lewis has a reason. He asserts: A rational person who envisages the different possible episodes that further experience may bring him will also know (if we abstract from practical limitations) what he will believe in each case, at the end. That seems to be equivalent, in this context, to the assertion that a rational person is committed beforehand to a recipe for belief-revision. How else could he know what he would come to believe after any possible course of further experience, except by being so committed?
>
> But must we accept this assertion about the rational person? What of the person who says: 'I can envision all of these possible episodes, one and only one of which will come to pass – I do not know now exactly what opinions and expectations I will form in response, I shall in most respects make up my mind then and there, hopeful and confident that I shall proceed both rationally and creatively' – is he not rational? It will be no use to object to him that if he expected to proceed rationally, then he would now expect to proceed by rules [KM: of which the only permitted one is conditionalization]. For that is the very point at issue, the point disputed. I conclude that rationality does not require conditionalization, nor does it require any commitment to follow a rule devised beforehand.

The idea, I take it, is that rational updating essentially involves the exercise of certain epistemic virtues, which are deployed in the agent's choice between two or more posterior credences on the assignment of an input value to her evidence. Of course only one of these posterior credences can result from conditionalization.

This line of thought also comes across in van Fraassen's derision of an epistemology which subscribes to "the idea that rationality would require a recipe-like rule" (2002, 144-145). Bayesian Updating, he says, relegates the rational thinker to being "just like Carnap's robot: his senses bring him propositions that he takes as evidence, and his total response to this consists in *conditionalizing* his present state of opinion on those propositions" (1984, 154). He characterizes the Bayesian as someone "who sees conditionalization – mere logical assimilation of the evidence – as the sole motive force in our epistemic life" (1989, 170). For emphasis he calls a model of evidential input which employs Jeffrey conditionalization the 'robot model' (1989, Chapter 13). Bayesian Updating allegedly straitjackets agents from exercising the epistemic virtues which rationality surely permits

them to exercise. The particular virtues which van Fraassen mentions by name are 'spontaneity', 'courage', and 'creativity' or 'innovation'.[182] These are virtues, he seems to think, which cannot be genuinely exercised in the choice of posterior credence without the agent's having a rational choice between two or more such credences.

The jist of van Fraassen's argument for Permissive Updating is that, if it were false (and hence if rationality were directive), it would not be rationally possible, once an agent has assigned an input credence to her total evidence, for her to exercise these virtues in her choice of posterior credence. If updating were not permissive – and more particularly, if it had to proceed by conditionalization – then rationality would run roughshod over the "rational free enterprise of the spirit" (1989, 172). Van Fraassen's target theory of updating is of course any which includes Bayesian Updating, which he takes to be directive (but which I'll argue shortly is not). But his argument could just as easily be targeted against any directive epistemology, and for now we'll just focus on his reasons for thinking that rationality permits the agent to choose a posterior credence by employing the sorts of virtues which cannot be genuinely exercised without there being two or more options. I'll argue that rationality requires no such thing, though we'll see later that Expansive Bayesianism is certainly compatible with the exercise of epistemic virtues, both in the agent's choice of update and in other facets of her credence formation.

First, however, van Fraassen's argument for Permissive Updating rests on the highly contentious assumption that agents have the psychological ability to choose what posterior to adopt on assigning an input value to their evidence at all. Chapter 4 has argued against doxastic voluntarism. But let's leave it aside for now and discuss the alleged epistemic virtues he appeals to.

Note first that spontaneity, creativity, and courage may be virtues in just the way van Fraassen describes when they are possessed by agents who are less than ideally rational. When one is dangling from the edge of a cliff, then perhaps some degree of courage, spontaneity, and creativity are needed to negotiate multiple options which offer some amelioration. In general, virtues can be exercised as correctives to inappropriate behavior: someone who has just put his foot in his mouth can manifest the virtue of graciousness by apologizing, even though he should never have put his foot in his mouth to begin with. The epistemic case might be similar: virtues might be operative in non-ideal cases in ways in which they are not operative in ideal cases. But like I am, van Fraassen is talking about ideal rationality, so life-saving or compensating moves will not be required.

He makes much of *courage*. Courage may be virtuous, but what is *epistemically* virtuous about it? It makes sense to think of epistemic courage as the virtue by which an agent opens himself up to evidential experiences which he could not anticipate, which is to say, about which he has no defined

[182] See also: "I declare innovation to be rational" (1989, 174), and his approval of "theoretical innovation and courageous embrace of new hypotheses that have gained my admiration" (*ibid*.)

credences at all: come what may, he is prepared to adopt even credences which he doesn't like. But this is not what van Fraassen thinks is courageous; indeed, unanticipated evidence does not feature at all in van Fraassen's view. Instead, van Fraassen ascribes courage to the agent who receives *exactly* the evidence he anticipated and, on receiving it, exercises permissive choice, here by choosing between two posterior credences to adopt on receipt of it.[183]

What might make such behavior courageous? Perhaps what van Fraassen has in mind is a certain 'courage of one's convictions': when the evidence comes in, the agent realizes that it makes some proposition *more* subjectively probable than he had anticipated beforehand. The thought would be that there is something about the evidential experience itself which conveys more information than the proposition to which it gives rise. But on the view of evidence articulated in Chapter 3, this is impossible: the content of the proposition (say, *e*) *just is* what the experience presents to the agent.[184] Any allegedly 'additional information' conveyed by an experience is evidence beyond *e*, which we may call *e**. If an agent receives a proposition *e** as evidence along with *e*, then van Fraassen is right that he does well not to conditionalize on *e* alone. But van Fraassen is wrong in suggesting that the answer is not to conditionalize; rather, it is to conditionalize on the conjunction of *e* and *e**. Thus on van Fraassen's view, the 'extra something' in the evidential experience really is something *non-evidential*. The courage which van Fraassen appeals to is a courage in departing from one's evidence. Van Fraassen, then, rejects clause (ii) of Evidentialism (see Chapter 1), that the receipt of new evidence is the only thing which rationally permits an update.

Arguing against van Fraassen from Evidentialism would thus be question-begging. Instead let's look at courage and the other alleged virtues in more detail. Such behavior – departing from one's evidence – may well be courageous (or foolhardy), but even if so, why should it be epistemically virtuous? It is not clear what might give such behavior a high expected truth-utility (or truth utility full stop),[185] and van Fraassen cannot simply say that it is epistemically virtuous because it is rational, since the claim that it is rational is precisely what I deny.

Another alleged virtue which van Fraassen lionizes is epistemic *spontaneity,* which an agent exercises by postponing her decision about what posterior to adopt until the moment in which she receives her evidence and assigns a credence to it. But what is valuable about this? Indeed, it makes sense to think it is epistemically worse to wait until evidence is received to decide what credences to adopt in light of it, for agents are arguably more impulsive and less discerning 'in the moment' and more prone to being swayed by non-evidential considerations. Shouldn't the evaluation of evidence

[183] We see this in (1989), which describes the rational and courageous agent who "can envision all of these possible episodes, one and only one of which will come to pass" (174).
[184] Van Fraassen, as far as I can tell, says nothing to suggest disagreement with my account of evidence.
[185] For discussion of truth-utility see Chapter 1.

best not be left, literally, to the 'spur of the moment'? I see little epistemic virtue in spontaneity, at least as van Fraassen conceives of it.

Van Fraassen also praises the epistemic *creativity* which voluntarist agents are free to exhibit in their choice of posterior credence. But what is epistemically virtuous about creativity? Credence formation is not about world-making, but rather world-representing, trying to be true to what is already there. Perhaps the idea is that the agent may be creative in forming a more accurate representation of the world (however we cash out 'accurate') by opting for one posterior rather than another. But, as above, it is hard to see what sense of 'better' might be at issue. There seems little reason to think that creativity issues in either higher truth utility or higher expected truth utility, and to say that it is *ipso facto* more rational than non-creativity is question-begging. I don't see what is epistemically desirable about creative choice of posterior credence.

In summary, the alleged virtues which van Fraassen adduces are not virtues at all, at least not as he construes them. But van Fraassen's virtues anyway presuppose doxastic voluntarism, which Chapter 4 argued is false. Moreover, we'll see below that, if any of them are virtues, Expansive Bayesianism can incorporate them – and without needing doxastic voluntarism to do so.

Complex Evidence Cases

A prima facie more promising route for motivating Permissive Updating is to appeal to our intuitions in cases in which agents deliberate between two or more posterior credences to adopt in the light of some total body of evidence. Such cases, which Chapter 4 called *complex evidence cases*, are characterized by an agent's acquiring a large or complicated quantity of evidence, much of which may be of a highly theoretical (as opposed to merely sensory) nature, and which does not at all clearly 'point' that agent toward one conclusion or another. Recall the agent from Chapter 4, who is the only juror on a criminal case and is charged to produce a subjective probability for the proposition Not-Guilty; and who, during the case, is faced with evidence which is extremely variegated, incohesive, complicated, technical, and detailed. When the agent interprets the evidence in one way, she is strongly inclined to a credence of x, but when she interprets it in another way, she feels just as strongly inclined to a different credence, of y. Turning her attention to one or the other interpretation puts her in a temporary state which resembles having the credence in question, except that, when she turns her attention to the other interpretation, her state changes. Finally, she winds up with a credence of x.

Chapter 4 defused an argument which appealed to cases such as this in the attempt to show that updating is *strongly voluntary*, in the sense that the agent has the psychological ability to choose

among two or more credences to assign on receipt of a given body of evidence. Even though such cases do not indicate that updating is strongly voluntary, they might be wielded to argue that updating is permissive. Such an argument would appeal to the intuition, which the case is apparently expected to provoke in readers, that the agent is rational in forming a credence of *x* in Not-Guilty but would have been equally rational had she assigned it a credence of *y* instead, even though she has no evidence discriminating between the two interpretations. And if she would be rational in both cases, then updating is permissive: there can be two or more alternative rational posterior credences to adopt on the receipt of total evidence E.

But why think that the complex-evidence agent would be rational to assign *x* rather than *y*, or vice-versa, even though she has no evidence discriminating between them? The complex evidence case may be regarded as epistemic analogue of a Buridan's Ass scenario in decision theory. In a Buridan's Ass scenario, an agent has the same expected utility assignment for each of two options, eating chocolate ice cream and eating vanilla ice cream. Surely, the thought would be, the rational agent can pick one, even though neither has a higher expected utility. Similarly, in the complex evidence case, the thought is that, just as the Buridan's ass agent can rationally pick chocolate or vanilla, the complex evidence agent can rationally update to a credence of *x* or of *y*. On this view, some instances of rational credence formation can be every bit as permissive as the most permissive of rational actions. And surely the complex evidence case is one such instance if ever there were one.

Let's concede that rational credence formation can be every bit as permissive as rational action. But it is a mistake to think that rational action is permissive at all: a (prudentially) rational agent acts so as to maximize expected utility. Hence, if there are no actions available to the Buridan's ass agent which have a higher expected utility than vanilla or chocolate, then choosing between them is *not* rational. Similarly, since there are no credences which the complex evidence agent can rationally form on the basis of her total evidence other than *x* or *y* which have a higher expected truth utility than these, she is not rational for adopting *x* over *y*.

Here the supporter of Permissive Updating might challenge us to account for the intuition that the Buridan's Ass agent would be rational for choosing either chocolate or vanilla. The response is that there are *similar* Buridan's Ass cases in which the agent would be rational for choosing chocolate, and would suggest that our intuitions about these cases might be driving any intuitions that the case as described is rational. For example, take what we'll call the *non-maximal Buridan's Ass case*. In this case, chocolate and vanilla have equal expected utilities, but not the highest. The agent assigns a higher expected utility to the choice of having some ice cream rather than none. She must merely solve the problem of which flavor to opt for in spite of her lack of reasons to prefer one

over the other. I suggest that she solves this problem, and comes out rational for it, by stipulating a prudential reason for choosing chocolate: for example, that a tossed coin lands heads. The result of the coin toss steps in as a "rookie" prudential reason. Allowing her decision to be determined by a coin toss will enable her to maximize expected utility by having some ice cream rather than none, and no permissive choice is involved at all.

Is there an analogous move which could show the complex-evidence agent to be rational in opting for either a credence of *x* or of *y*? There is not. The complex-evidence agent cannot rationally stipulate a reason for opting for *x* or for *y*, unless that reason constitutes evidence, which is the only sort of reason for changing credence. In this case, evidence would pertain to whether the defendant is guilty. There is only one possible scenario in which a stipulated reason, such as the result of a coin toss, could provide evidence: one in which the proposition about which the agent's credences are wavering pertains to the outcome of the coin toss itself. But here, flipping a coin amounts to collecting more evidence. Since the complex evidence case is premised on the agent's evidence being indecisive, any reason she chooses for adopting one credence over another would have to be one which does not raise or lower her probability and hence does not constitute evidence.

Here is a suggested diagnosis of why the Buridan's Ass analogy might tempt the supporter of Permissive Updating to think that it supports her view. Sometimes, although an agent may be torn between two credences to form on the basis of a body of total evidence, she must *act as if* one credence has a higher expected truth utility than another. The scientist must pursue one theory or another (on pain of losing his grant money), a doctor must follow some course of treatment or other (lest the patient deteriorate rapidly), a Pascalian agnostic must go to church or not (since these exhaust the logical options). If one thinks that credence can be measured by an agent's actions, then one might be tempted to conclude that such actions do indicate the credences of the agents performing them. But I have rejected the view that credences are *defined* in terms of such measurement: actions may indicate credences, but not infallibly.

Bayesian Updating is Compatible with Permissive Updating: Expansive Bayesianism

None of the considerations we have advanced for Permissive Updating support it particularly well if at all. Permissive Updating is not a promising premise on which to stake the fate of Bayesian Updating, or indeed anything else.

But even if Permissive Updating does happen to be true, the argument from it to the negation of Bayesian Updating is unsound. Premise 2 is false: Bayesian Updating does not entail the negation of Permissive Updating. Bayesianism is consistent with the claim that rationality is sometimes

compatible with the assignment of more than one posterior credence on receipt of any given piece of evidence.

How can this be, if Bayesianism requires that updates only ever proceed by conditionalization? If anything would seem to be directive rather than permissive, it would be Bayesian Updating. But Expansive Bayesianism has scope for at least two ways in which agents might – assuming that Permissive Updating is true – have multiple options for her posterior credence. Recall my claim that an agent may acquire conditional credences when he had none before. I suggested that this may occur when the agent acquires evidence in which he did not, before the input step, have a non-zero credence. Such evidence, we saw, comes in two varieties. One is *impossible evidence*, evidence in which, before the input step, the agent's credence was zero. The other is *undefined evidence*, which the agent did not grasp before the evidential experience and hence given which she had no conditional credences before this. I argued that evidential experiences may produce such evidence and that, when they do, they provide the occasion for the agent to form completely new conditional credences; namely, credences conditional on that evidence. If this is so, then, *pace* van Fraassen, there is at least potentially room for rational options vis-à-vis conditional credences: when such credences are formed conditional on impossible or undefined evidence.

The idea, which I neither endorse nor reject but merely suggest to show that Bayesianism is compatible with some degree of permissiveness, would be something like this. Even though Bayesianism does not permit the agent, without evidence, to change the conditional credences she already has, it has no provisions for restricting the conditional credences which she acquires on receipt of undefined or impossible evidence. Rational updating might be directive, prescribing the precise value of the conditional credences she adopts, and hence of the posterior she ought rationally to assign given the input value assigned to her evidence. Or rational updating may be permissive, permitting her the option of two or more posteriors given that input value. Both claims are consistent with Bayesian updating.

One might think that impossible- and undefined-evidence scenarios are uncommon, which would mean that, if agents are rationally permitted multiple conditional credences, such cases will be rare. This result might contravene the ethos, if not the letter, of Permissive Updating. But on closer inspection there is reason to think that such scenarios might occur rather frequently. More specifically, there is reason to think that *undefined-evidence* scenarios might occur more frequently than one might think, even if impossible-evidence scenarios remain rare. Recall the argument in Chapter 3 for the claim:

Coming to grasp new propositions is a common feature of agents' epistemic lives. (*Common*)

First, it happens, I argued, whenever agents grasp new theoretical claims, whether they are schoolchildren, undergraduates, or seasoned professors. Second, it happens whenever agents have even somewhat complex perceptual experiences, from the utterly unprecedented to the merely overloading. If undefined-evidence scenarios occur as frequently as these rather quotidian experiences do, then permissive choice of conditional credence, and hence of the posterior an agent adopts when she assigns an input value to undefined evidence, could rationally occur rather often, if it can rationally occur at all.

Let us return to our discussion of van Fraassen's epistemic virtues. Expansive Bayesianism allows for the exercise (or failure to exercise) of significant epistemic virtue, and hence for a great deal of the epistemic agency so important to van Fraassen. Moreover, Bayesian Updating allows these virtues to be exercised in ways which are more obviously virtuous than they are on van Fraassen's view. Take courage. It makes sense to think of epistemic courage as the virtue by which an agent opens himself up to evidential experiences which he could not anticipate in the hope of acquiring undefined or even impossible evidence. The courage lies in doing this and in bracing oneself to form even credences which surprise or dismay one. As for spontaneity, I suggested above that it is no epistemic virtue at all. But even if it were, it could be exercised in the spontaneity with which one seeks out experiences which might provide undefined or impossible evidence. Finally, as for creativity: I argued that it hard to see how creativity in rejecting or choosing one's posterior credence could be an epistemic virtue. But there may be virtue in the creativity with which one seeks undefined or even impossible evidence and hence puts oneself in a position to gain new conditional credences the values of which one may perhaps also exercise virtuous agency in setting. To return briefly to Chapter 4's argument against strong doxastic voluntarism, note that the exercise of none of these virtues or alleged virtues, on an Expansive Bayesian framework, requires the exercise of strongly voluntary choice of posterior credence.

As will now be clear, Expansive Bayesianism shows how agents can rationally resolve complex evidence cases, in which they are wavering between two or more posterior credences to assign to a proposition given some input value for their total evidence: they can seek out undefined- or even impossible-evidence scenarios to provide them with evidence.

Bayesian Updating, then, does not entail the denial of Permissive Updating, but nor does it entail Permissive Updating. It leaves room for permissive updating should such room be needed, since there will always be more undefined evidence, and perhaps even impossible evidence, to acquire.

However, there is a stronger claim with which Bayesian Updating is incompatible:

All rational updating is permissive. (*Strong Permissive Updating*)

Permissive Updating says that *some* rational updates are permissive, and Expansive Bayesianism gives scope for agreeing: precisely those updates may be permissive which occur on the basis of undefined or impossible evidence. But updates in all other cases, says Bayesian Updating, are directive.

Thus one could argue to the falsehood of Bayesian Updating by providing strong reasons to think that all rational updates are permissive. However, given the difficulties in establishing even that some are, I doubt that a successful case can be made for the stronger claim. Moreover, there is reason to think that Strong Permissive Updating is too extreme. Strong Permissive Updating says in effect that rationality is never satisfied *unless* there are at least two rational posterior credences which one may adopt on assigning an input value to one's evidence. But if rationality required the existence of multiple options, it would be unique among sources of normativity for doing so. Take goodness for example. Agents are perfectly able to act rightly even in situations where there is only one action which morality permits them to perform; they display it by performing that action instead of some wrong action. An agent who rescues a kitten when there was no other morally right choice in the situation, has done the right thing regardless of whether she could instead have chosen to rescue a puppy.[186] A similar point holds for the norm of health: when there is only cake and Brussels sprouts on offer, health requires that the agent choose Brussels sprouts. Health does not also require that there *also* have been another healthy option on offer, such as quinoa, in order for the choice of Brussels sprouts to be healthy. The existence of other healthy choices is irrelevant. I do not see why rationality should be such that there are *always* two or more rational posterior credences to adopt in response to an evidential input value.

Expansive Bayesianism has room for any permissiveness which there may be, and no more.

Conclusion: Room for a Different Form of Permissiveness: The Input Step

Bayesian Updating has room for a milder form of permissiveness which applies even in updating scenarios which do not involve undefined- or impossible-evidence scenarios. Recall that the kind of permissiveness in question up till now has pertained to different posterior credences which one might rationally adopt once one has assigned an input value to one's evidence. A slightly milder form

[186] Let's assume that neither action was better either intrinsically or due to circumstantial features.

of permissiveness, which Bayesian Updating also makes room for, pertains to rational options among a variety of input values themselves. Since the agent's input value has a role in determining the value of her posterior credence in a proposition *p*, the assigning of one input value rather than another will affect the agent's posterior credence in *p*. Thus Bayesian Updating has room for this milder form of permissiveness, too, which we'll characterize as the claim that:

> At least some input values for one's evidence are permissive. (*Permissive Input*)[187]

Of course, agents do not choose the experiences they have (although of course they can act in ways that influence them), but there may be epistemic virtues at play in how one construes them. (The exercise of such virtue need not be voluntary: rather, agents can over time cultivate virtuous epistemic dispositions which will manifest themselves involuntarily in the appropriate circumstances.) Of course, this does not mean that Permissive Input is *true* if Bayesian Updating is. Rationality might not permit agents to exercise choice about which credence they assign to their evidence proposition.

Why might one think that the input step rationally allows for multiple input values? The reason is this: the input step occurs in response to an evidential experience. To say that there is no scope for permissive input in response in response to an evidential experience is to say that,

> For any evidential experience, there is only *one* credence which an agent may rationally assign to the proposition to which that experience gives rise. (*Unique Input*)

Unique Input is a strong claim indeed, placing very strict rational constraints on an agent's unmediated doxastic response to experience. It would have to be supported by a correspondingly strong argument, of which there is presently none of which I am aware.

One might argue for Unique Input by making some sort of appeal to the proper function of human beings *qua* epistemic agents. That is, given the kind of being we are, the environment in which our cognitive functions are designed (say, by God or by evolution[188]) to operate, and the 'design plan' which specifies how we ought rationally to respond to any given input, there are unique values which we rationally ought to assign to the propositions to which evidential experiences give rise. An argument along these lines might have some merit in supporting the claim that there are *some* rational restrictions on the credence an agent may assign to her evidence proposition. But such an argument would have to do much more in order to show that there is *one*

[187] This is an example of a rational constraint on the value of an agent's input credence, discussed in Chapter 3, Part II.
[188] See e.g. Millikan (1987), Plantinga (2000).

and only one appropriate response to one's evidence. The proponent of such an argument would be hard pressed, for one, to say what that unique value should be for any given experience or type of experience, or even to say why there should be *one* unique value rather than two or more (which would be enough to secure Permissive Updating). But to my knowledge no such argument has been offered for Unique Input, and I am not about to try.

Thus Bayesian Updating, by incorporating an input step in every update, secures theoretical room for a form of permissiveness in assigning a credence to one's evidence and hence, indirectly, to the posterior one winds up adopting after updating on that evidence. This is just another way in which Bayesianism is compatible with a good degree of epistemic permissiveness.

Bayesian Updating is a much more flexible theory of rational updating than liberationist critics give it credit for.

References

Alston, William (1988). 'An Internalist Externalism', in *Synthese* 74: 265-283.

Armendt, Brad (1980). 'Is there a Dutch Book Argument for Probability Kinematics? *Philosophy of Science*, vol. 47: 583-589.

Bach, Kent (1985). 'A Rationale for Reliabilism', *The Monist* 68: 246-263.

Bernecker, Sven and Fred Dretske (2000) (eds.). *Readings in Contemporary Epistemology*, Oxford: Oxford University Press.

Brentano, Franz (1874/1911/1973). *Psychology from an Empirical Standpoint*, London: Routledge and Kegan Paul.

Brewer, Bill (1999). *Perception and Reason*, Oxford: Oxford University Press.

Brewer, W. F. and B.L. Lambert (2001). 'The Theory-Ladenness of Observation and the Theory-Ladenness of the Rest of the Scientific Process', *Philosophy of Science*, 68(3): 176-186.

Brown, Harold I. (1983). 'Incommensurability', in *Inquiry* 26 (1): 3 – 29.

Chalmers, David J. (2011). 'Frege's Puzzle and the Objects of Credence', in *Mind* 120 (479):587-635.

Chisholm, Roderick (1977). *Theory of Knowledge* , 2nd Edition, Englewood Cliffs: Prentice-Hall.

Christensen, David (1992). 'Confirmational Holism and Bayesian Epistemology', in *Philosophy of Science*, 59: 540-57.

Christensen, David (2004). *Putting Logic in its Place*, Oxford: Oxford University Press.

Christensen, David (2007). 'Epistemology of Disagreement: The Good News', in *Philosophical Review* 116: 187-218.

Christensen, David (2007). Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals', in *Oxford Studies in Epistemology*, Oxford: Oxford University Press, 2: 3-31.

Clifford, W. K. (1879). *Lectures and Essays*, London: Macmillan and Co.

Conee, Earl and Richard Feldman (2001). 'Internalism Defended', in Hilary Kornblith, ed., *Epistemology: Internalism and Externalism*, Cambridge: MIT Press.

Conee, Earl, and Richard Feldman (2004). *Evidentialism*, Oxford: Oxford University Press.
de Finetti, Bruno (1990): *Theory of Probability*, Vol. I (first edition 1974), New York: Wiley.

Descartes, René (1641). 'Meditations on First Philosophy', in *The Philosophical Writings of Descartes*, vol. II, trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch, Cambridge: Cambridge University Press, 1984.

Domotor, Zoltan (1980). 'Probability Kinematics and Representation of Belief Change', *Philosophy of Science*, vol. 47: 384-403.

Döring, Frank (1999). 'Why Bayesian Psychology is Incomplete', in *Philosophy of Science* 66 (Proceedings): S379–S389.

Dougherty, Trent (2011). 'In Defense of Propositionalism about Evidence', in *Evidentialism and its Discontents*, Oxford: Oxford University Press, Trent Dougherty, ed., 226-231.

Dretske, Fred (1970). 'Epistemic Operators', *Journal of Philosophy* 67 (24): 1007–1023.

Dunn, Jeffrey (2010). *Bayesianism and Having Evidence*, PhD Dissertation, University of Massachusetts at Amherst, Open Access Dissertations: http://scholarworks.umass.edu/open_access_dissertations/273/

Eagle, Antony (2011), ed. *Philosophy of Probability: Contemporary Readings*, London and New York: Routledge.

Earman, John (1992). *Bayes or Bust?* Cambridge, Massachusetts: MIT Press.

Elga, Adam (2000). 'Self-Locating Belief and the Sleeping Beauty Problem', *Analysis* 60, 143-147.

Elga, Adam (2007). 'Reflection and Disagreement', in *Noûs*, 41: 478-502.

Elga, Adam (2010). '*Subjective Probabilities Should be Sharp*', in *Philosopher's Imprint*, vol. 10, no. 5.

Eriksson, Lina and Alan Hájek (2007). 'What are Degrees of Belief?' *Studia Logica* 86 (2): 185-215

Feldman, Richard (2007). 'Reasonable Religious Disagreements', in *Philosophers Without Gods: Meditations on Atheism and the Secular Life*, Louise Antony, ed., Oxford: Oxford University Press: 194-214.

Feldman, Richard and Ted A. Warfield (2010), eds. *Disagreement*. Oxford: Oxford University Press.

Feyerabend, Paul (1975). *Against Method: Outline of an Anarchistic Theory of Knowledge*, London: New Left Books.

Field, Hartry (1978). 'A Note on Jeffrey Conditionalization', in *Philosophy of Science* 45 (3) 361-367.

Foley, Richard (1993). *Working Without a Net: A Study of Egocentric Epistemology*, New York: Oxford University Press.

Foley, Richard (2009). 'Beliefs, Degrees of Belief, and the Lockean Thesis', in *Degrees of Belief*, Franz Huber and Christoph Schmidt-Petri, eds., Springer.

Frankish, Keith (2009). 'Partial Belief and Flat-Out Belief', in *Degrees of Belief*, Franz Huber and Christoph Schmidt-Petri, eds., Springer.

Gabbay, Dov M.; Stephan Woods; and John Hartmann (2011). *Handbook of the History of Logic, Volume 10: Inductive Logic*, Amsterdam: North Holland.

Garber, Daniel (1980). 'Field and Jeffrey Conditionalization', in *Philosophy of Science* 47: 142-145.

Garber, Daniel (1983). 'Old Evidence and Logical Omniscience in Bayesian Confirmation Theory', in John Earman (ed.), *Testing Scientific Theories*, Minneapolis: University of Minnesota Press.

Ginet, Carl (1975). *Knowledge, Perception, and Memory*, Dordrecht: D. Reidel.

Goldman, Alvin (2001). 'Internalism Exposed', in Hilary Kornblith, ed., *Epistemology: Internalism and Externalism*, Cambridge, Massachusetts: MIT Press.

Greaves, Hilary and David Wallace (2006). 'Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility, *Mind* 115: 607-632.

Gunther, York (2003) (ed.). Essays on Nonconceptual Content, Cambridge, MA: MIT Press.
Haenni, Rolf (2009). 'Non-Additive Degrees of Belief', in Frans Huber and Christoph Schmidt-Petri, eds., *Degrees of Belief*, Springer.

Hájek, Alan (2003). 'What Conditional Probabilities Could Not Be', in *Synthese* 137 (3), 273-323.

Hájek, Alan (2008). 'Arguments for – or against – Probabilism?' in *British Journal for the Philosophy of Science* 59: 793-819.

Hawthorne, James (2009). 'The Lockean Thesis and the Logic of Belief', in *Degrees of Belief*, Franz Huber and Christoph Schmidt-Petri, eds., Springer.

Horwich, Paul (1992). 'Bayesian Problem-Solving and the Dispensibility of Truth', in *Philosophical Issues* 2: 205-214.

Howson, Colin (2000). *Hume's Problem*, Oxford: Oxford University Press.
Howson, Colin, and Peter Urbach (1993). *Scientific Reasoning: The Bayesian Approach*, 2nd Edition, Chicago and La Salle, Illinois: Open Court.

Hoyningen-Huene, P. (1993). *Reconstructing Scientific Revolutions: The Philosophy of Science of Thomas S. Kuhn*, Chicago: University of Chicago Press.

Huber, Franz, and Christoph Schmidt-Petri (2009), eds. *Degrees of Belief*, Springer.

Hume, David (1975). *A Treatise of Human Nature*, L. A. Selby-Bigge, ed., 2nd ed. revised by P. H. Nidditch, Oxford: Clarendon Press.

Husserl, Edmund (1900/1970). *Logical Investigations*, J. N. Findlay, Translator, London: Routledge and Kegan Paul.

James, William (1956). 'The Will to Believe', in *Will to Believe and Human Immortality*, New York: Dover Publications.

James, William (1956). 'The Sentiment of Rationality', in *Will to Believe and Human Immortality*, New York: Dover Publications.

Jeffrey, Richard (1983). *The Logic of Decision*, 2nd ed., Chicago: University of Chicago Press.

Jeffrey, Richard (1992). *Probability and the Art of Judgment*. Cambridge: Cambridge University Press.

Jeffrey, Richard (2004). *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press.

Joyce, James (1998). 'A Non-Pragmatic Vindication of Probabilism', in *Philosophy of Science* 65, 575-603.

Joyce, James (2009). 'Accuracy and Coherence', in *Degrees of Belief*, Franz Huber and Christoph Schmidt-Petri, Eds., Springer.

Kelly, Thomas (2003). 'Epistemic Rationality as Instrumental Rationality: A Critique', in *Philosophy and Phenomenological Research*, Vol. 66, No. 3: 612-640.

Kelly, Thomas (2005). 'The Epistemic Significance of Disagreement' in Tamar Szabó Gendler and John Hawthorne, eds., *Oxford Studies in Epistemology*, Vol. 1, Oxford: Oxford University Press.

Kierkegaard, Søren (1846 / 1992). *Concluding Unscientific Postscript to Philosophical Fragments*, Howard V. Hong and Edna H. Hong (trans.), Princeton: Princeton University Press.

Kornblith, Hilary (2001), ed. *Epistemology: Internalism and Externalism*, Cambridge, Massachusetts: MIT Press.

Kosso, P. (1992). *Reading the Book of Nature*, Cambridge: Cambridge University Press.

Kuhn, Thomas (1962). *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Lackey, Jennifer (2008). 'What Should We Do When We Disagree?' in In Tamar Szabó Gendler & John Hawthorne, eds., *Oxford Studies in Epistemology*, Oxford: Oxford University Press.

Lakatos, I. (1978). *The Methodology of Scientific Research Programs*, Cambridge: University Press.

Lange, Marc (1999). 'Calibration and the Epistemological Role of Bayesian Conditionalization', *Journal of Philosophy*, Vol. 96: 294-324.

Lange, Marc (2000). 'Is Jeffrey Conditionalization Defective by Virtue of Being Noncommutative? Remarks on the Sameness of Sensory Experience', in *Synthese*, 123(3): 393–403.

Levi, Isaac (1967). 'Probability Kinematics', in *British Journal for the Philosophy of Science*, vol. 18, 197-209.

Lewis, C. I. (1929). *Mind and the World Order*, Toronto, Ontario: General Publishing Company, Ltd.

Lewis, C. I. (1947). *An Analysis of Knowledge and Valuation*, LaSalle, Illinois: Open Court.

Lewis, David (1980). 'A Subjectivist's Guide to Objective Chance', in *Studies in Inductive Logic and Probability 2*, Richard. C. Jeffrey, ed., Berkeley: University of California Press.

Lewis, David (1983). 'Attitudes *De Dicto* and *De Se*', in *Philosophical Papers*, vol. 1, Oxford: Oxford University Press: 133-159.

Lewis, David (1999). 'Why Conditionalize?' in *Papers in Metaphysics and Epistemology* Volume 2. Cambridge: Cambridge University Press, 403-407.

Mackie, John (1982). *The Miracle of Theism: Arguments for and Against the Existence of God*, Oxford: Oxford University Press.

Maher, Patrick (1993). *Betting on Theories*, Cambridge: Cambridge University Press.

Maher, Patrick. (1995). 'Probabilities for New Theories', in *Philosophical Studies*, vol. 77, no. 1: 103-115.

Meacham, Christopher J. G. (2008). 'Sleeping Beauty and the Dynamics of De Se Beliefs', in *Philosophical Studies*, 138: 245-269.

Meacham, Christopher J. G. (2010). 'Unravelling the Tangled Web: Continuity, Internalism, Uniqueness and Self-Locating Belief', in *Oxford Studies in Epistemology*, Vol. 3, Tamar Szabó Gendler and John Hawthorne, eds., Oxford: Oxford University Press.

McDowell, J. H. (1994). *Mind and World*, Cambriedge, Massachusetts: Harvard University Press.

Millikan, Ruth Garrett (1987). *Language, Thought, and Other Biological Categories: New Foundations for Realism*, Cambridge: MIT Press.

Milne, P. (2003). 'Bayesianism vs. Scientific Realism', *Analysis*, 63, pp. 281-288.

Monton, Bradley (2007), ed. *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas C. van Fraassen*, Oxford: Oxford University Press.

Olsson, Erik J. and Stefan Schubert (2007). 'Reliability Conducive Measures of Coherence', in *Synthese* 157 (3): 297 - 308.

Peacocke, Christopher (1992). Cambridge, Massachusetts: MIT Press.

Plantinga, Alvin (1993). *Warrant and Proper Function*, Oxford: Oxford University Press.

Plantinga, Alvin (2000). *Warranted Christian Belief*, Oxford: Oxford University Press.

Plato (1961). *Meno*, in *The Collected Dialogues*, Edith Hamilton and Huntington Cairns (eds.), Princeton: Princeton University Press.

Pojman, L. (1999) (ed.). *The Theory of Knowledge: Classical and Contemporary Readings*, 2nd edition, Belmont: Wadsworth Publishing.

Popper, K. R. (1963). *Conjectures and Refutations*, New York: Harper and Row.

Psillos, Stathis (2007). 'Putting a Bridle on Irrationality: An Appraisal of van Fraassen's New Epistemology', in *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas C. van Fraassen*, Bradley Monton, ed., Oxford: Oxford University Press, 134-164.

Quine, W. V. O. (1969). 'Naturalized Epistemology', in *Ontological Relativity and Other Essays*, New York: Columbia University Press, 69-90, re-printed in Sven Bernecker and
Fred Dretske (eds.), *Readings in Contemporary Epistemology*, Oxford: Oxford University Press, 2000.

Ramsey, Frank P. (1931). 'Truth and Probability', in *The Foundations of Mathematics and Other Logical Essays*, R. B. Braithwaite, ed., London: Routledge and Kegan Paul: 156-198.

Russell, Bertrand (1917 / 1951). 'The Relation of Sense-Data to Physics', in *Mysticism and Logic,* London: George Allen & Unwin Ltd., reprinted Totowa, New Jersey: Barnes & Noble Books.

Sellars, Wilfrid (1956). 'Empiricism and the Philosophy of Mind', in *Minnesota Studies in the Philosophy of Science*, vol. I, H. Feigl and M. Scriven (eds.), Minneapolis, Minnesota: University of Minnesota Press: 253–329.

Shimony, Abner (1970). 'Scientific Inference', in R. G. Colodny (ed.), *Pittsburgh Studies in the Philosophy of Science*, vol. 4.

Shimony, Abner (1985). 'The Status of the Principle of Maximum Entropy', in *Synthese* 63, 35-53.

Skyrms, Brian (1987). 'Dynamic coherence and Probability Kinematics', in *Philosophy of Science*, vol. 54: 1-20.

Stalnaker, Robert C. (2008). *Our Knowledge of the Internal World*, Oxford: Oxford University Press.

Steup, Matthias (1999). 'A Defense of Internalism', in L. Pojman (ed), *The Theory of Knowledge: Classical and Contemporary Readings*, 2nd edition, Belmont: Wadsworth Publishing.

Sturgeon, Scott (2008). 'Reason and the Grain of Belief', in *Noûs*, 42(1): 139-165.

Sturgeon, Scott (2011). 'Confidence and Coarse-Grained Attitudes', in *Oxford Studies in Epistemology*, Vol 3: 126-149.

Suppes, Patrick (2007). 'Where do Bayesian Priors Come From?' in *Synthese* 156: 441-471.

Swinburne, Richard (2004). *The Existence of God*, Oxford: Clarendon Press.

Teller, Paul (1973). 'Conditionalization and Observation', in *Synthese*, vol. 26: 218-258.

van Fraassen, Bas (1984). 'Belief and the Will', *Journal of Philosophy* 81: 235-256, reprinted in Antony Eagle (2011), ed. *Philosophy of Probability: Contemporary Readings*, London and New York: Routledge. Page references are from the latter.

van Fraassen, Bas (1986). 'A Demonstration of the Jeffrey Conditionalization Rule', in *Erkenntnis* 24 (1): 17-24.

van Fraassen, Bas (1989). *Laws and Symmetry*, Oxford: Clarendon Press.

van Fraassen, Bas (1995). 'Belief and the Problem of Ulysses and the Sirens', in *Philosophical Studies*, vol. 77: 7-37.

van Fraassen, Bas (2002). *The Empirical Stance*, New Haven and London: Yale University Press.

White, Roger (2005). 'Epistemic Permissiveness', in *Philosophical Perspectives* 19 (1): 445–459.

White, Roger (2009). 'Evidential Symmetry and Mushy Credence', in *Oxford Studies in Epistemology*, Oxford: Oxford Univesity Press.

Williams, Bernard (1973). *Problems of the Self*, Cambridge: Cambridge University Press.

Williams, Robbie (forthcoming). 'Generalized Probabilism', in *Journal of Philosophical Logic*.

Williamson, Jon (2009). 'Philosophies of Probability', in *Philosophy of Mathematics (Handbook of the Philosophy of Science, Vol. 4)*, Gabbay, Dov, Paul Thaggard, and John Woods, eds., Amsterdam: Elsevier, 493-534.

Williamson, Timothy (2000). *Knowledge and Its Limits*, Oxford: Oxford University Press.

Wolterstorff, Nicholas (1996). *John Locke and the Ethics of Belief*, Cambridge: Cambridge University Press.