PROBABILISTIC ACCOUNTS OF INFERENTIAL JUSTIFICATION: LIBERALISM AND INFERENCE TO THE BEST EXPLANATION

by

Gregory E. Gates

B.S. in Mathematics, Indiana University, 2001

B.A. in Philosophy, Indiana University, 2001

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Gregory E. Gates

It was defended on

August 9, 2011

and approved by

Cian Dorr, Senior Faculty Lecturer, University of Oxford

Anil Gupta, Distinguished Professor

Mark Wilson, Professor

James Allen, Professor

Dissertation Advisors: Cian Dorr, Senior Faculty Lecturer, University of Oxford,

Anil Gupta, Distinguished Professor

PROBABILISTIC ACCOUNTS OF INFERENTIAL JUSTIFICATION: LIBERALISM AND INFERENCE TO THE BEST EXPLANATION

Gregory E. Gates, PhD

University of Pittsburgh, 2011

I argue for three main conclusions. First, we should adopt a "probability first" approach to epistemology, which takes facts about justification for outright belief to supervene on facts about rationally permissible credence distributions. Such an approach is plausible even though standard accounts that reduce belief to credence above a threshold or invariance in conditional preferences are vulnerable to intuitive counterexamples.

Second, I argue that a dogmatist response to skepticism about inferential justification is false if we adopt a probability first approach. Dogmatists hold that we might gain justification to believe $E \supset H$ for the first time when we learn E (and nothing stronger). I show that only a dynamic Keynesian model is compatible with dogmatism about inferential justification. But the main virtue of the dynamic Keynesian model—it allows for learning about fundamental evidential relationships—is by no means unique to it. I conclude that a rationalist liberalism, which holds that we are independently justified in believing $E \supset H$ whenever we're inferentially justified in believing H on the basis of E, is the best anti-skeptical account of inferential justification on most probabilistic models.

Finally, I argue that the compatibilist approach to the conflict between Bayesian conditionalization and inference to the best explanation (IBE) fails. However, we anyway need to impose constraints on rational credence other than conditionalization, and we should take explanatory considerations to constrain the rationally permissible prior credence distributions. I present an account of IBE such that we should give higher conditional prior credence to H, given E, when H is the most intellectually satisfying explanation of E, and defend

this account against the objection that the subjectivity of intellectual satisfaction will lead to an unacceptably permissive epistemology.

TABLE OF CONTENTS

LIST OF FIGURES

1.0 PROBABILISTIC EPISTEMOLOGY

1.1 OVERVIEW

I once intended this dissertation to defend the view that we have substantive a priori knowledge of contingent facts about evidential relationships. After delving deep into Bayesian epistemology, so as to understand the alleged incompatability of Bayesianism with dogmatism about perceptual justification and with inference to the best explanation, I have adopted a thoroughly probabilistic orientation to the questions that were once my main concern. While I am confident that sensible questions posed in the language of traditional epistemology (e.g., do we have substantive a priori knowledge of contingent facts about evidential relationships?) can be answered and ultimately must be answered using Bayesian tools, the conclusion I am willing to defend in this dissertation is now more limited. My primary claim is that, when an ideally rational subject S finds H to be the most intellectually satisfying explanation of some evidence E, given background assumptions K, she epistemically ought to have hypothetical prior credences $cr_\emptyset$ such that $cr_\emptyset(E \supset H|K)$ is relatively high. While this may not suffice in all cases to justify outright belief in $E \supset H$, or to make it the case that S knows that $E \supset H$, justified high credence is at least one necessary condition for justified belief.

Furthermore, if S is justified in having priors such that $cr_\emptyset(E \supset H|K)$ is relatively high, and if it is possible for S to gain justification for H, given background assumptions K, by learning E (and nothing stronger), then I claim that it is possible that S is conditionally justified in believing $E \supset H$, given K, in the absence of any propositional evidence gained from experience. This is, I take it, something *like* the claim that S is possibly justified in believing $E \supset H$ a priori, but I do not intend to defend the claim that defenders of a priori justification would find this a satisfying account of what they would mean by a priori justification to believe $E \supset H$.

I do not regard the scaling back of my original ambitions as a loss. For the reason one might hope that we have a priori knowledge of contingent evidential relationships—or anyway, the reason I wanted to defend that claim—is the worry that such knowledge is necessary to avoid skepticism about inferential justification. But on the probabilistic approach that I now prefer, if we assume that an ideally rational agent should have the credences that she would get by conditionalizing her hypothetical priors on her total evidence, then requiring that she give $E \supset H$ high conditional prior credence, given K, will force her to have high posterior credence in H, given EK. That is, high conditional prior credence in $E \supset H$ given K is correlated with having credences such that E confirms H, given K. So the skeptical worry that we are not justified in taking some evidence E to support H can be answered, in a Bayesian framework, so long as we permissibly have the prior credences that I claim we should. We do not need to suppose anything as strong as a priori justification to believe that E is evidence for H.

So I claim that we can be inferentially justified in believing H, on the basis of E, so long as H is an intellectually satisfying explanation of E. The defense of this view will have three main parts.

First, in this chapter, I defend my overall probabilistic approach. At the level of outright belief, there are three doxastic attitudes that exhaust the options with respect to any proposition we consider: belief, disbelief, or suspension of judgment. Facts about which of these attitudes it is permissible for an individual subject, at a time, to have to a proposition P, supervene on facts about which hypothetical prior credence distributions it is permissible for the subject to have at that time and her total evidence at that time. While I do not have a completely decisive argument in favor of this "probability first" approach to epistemological questions, I will argue that it is more plausible than some prominent views of the relationship between credence and belief make it seem. In particular, I argue that belief is not reducible to credence above a threshold, and that belief is not reducible to invariance in conditional preferences. I then offer a diagnosis of the reasons these reductive accounts fail: they attempt to make a subject's outright beliefs supervene only on the precise credence that the balance of her evidence supports (plus, perhaps, her preferences). I will show that considering other aspects of a subject's credence distribution allows us to distinguish between the cases that provide an intuitive counterexample to the overly simple views I reject. I conclude that justified belief plausibly supervenes on justified credence and the evidence, plus, possibly, other factors of her psychology that are not epistemically evaluable (such as preferences).

In chapter 2, I present a schematic skeptical argument that can be applied to individual cases of (alleged) inferential justification. When I claim that I am justified in believing H on the basis of some evidence E, the skeptic may attempt to show that I am not independently justified in believing $E \supset H$, and therefore my claim must be false. I then consider responses to this skeptical argument; my main interest is in the dogmatist response to this kind of skepticism which denies the need for independent justification to believe $E \supset H$ in such cases. White (2006) and Weatherson (2007) have claimed that, from Bayesian assumptions, it can be shown that dogmatism is false—I call this the Bayesian argument against dogmatism. Given the complexity of the relationship between outright belief and credence, discussed in chapter 1, this argument requires a linking principle to connect the claims it makes at the probabilistic level to the claims it makes at the level of outright belief. Neither White (2006) nor Weatherson (2007) provides a satisfying linking principle. I offer a linking principle that can fill in this gap in the Bayesian argument against dogmatism. I conclude that dogmatism is incompatible with traditional Bayesian assumptions. While Weatherson (2007) offers a generalization of the orthodox Bayesian model that is formally compatible with dogmatism, it is poorly motivated independently of a desire to defend dogmatism. Therefore, I urge that epistemologists who accept my probabilistic orientation should reject dogmatism in favor of rationalism. That is, probabilistic anti-skeptics should endorse the view that we do have independent justification to believe $E \supset H$ whenever we can gain inferential justification to believe H on the basis of some evidence E.

Finally, in chapter 3, I offer an account of the source of this independent justification. I claim that an ideally rational inquirer is justified in giving relatively high credence to $E \supset H$ when H is an intellectually satisfying explanation of E given her background knowledge. I defend this as the best Bayesian account of the role of inference to the best explanation in inquiry. Van Fraassen (1989) argues that inference to the best explanation is incompatible with Bayesian conditionalization, which is the sole norm of inferential justification. I agree with the first part of this claim (the incompatability) but not the second part (the endorsement of conditionalization). I agree with Weisberg (2009) that explanatory considerations constrain the permissible prior probability distributions. I then offer an account of explanatory value along the lines of White (2005b), according to which a hypothesis has explanatory value when it satisfies our curiosity. I anticipate that the main objection to this account of explanatory value will be that it is too subjective and will lead to an unacceptably permissive epistemology, according to which multiple posterior credence distributions are permissible and it may be impossible to reach agreement about what the evidence supports. I defend my view against this objection by appealing to hopefulness of IBE. It is self-correcting; there are limits on what ideally rational inquirers will find satisfying; and if we treat the intellectual satisfaction of our epistemic peers as relevant evidence that bears on $E \supset H$, we can reasonably expect that our evaluations of the force of the evidence will converge.

1.2 TRADITIONAL AND BAYESIAN EPISTEMOLOGY

Bayesian epistemologists, in the broadest sense, are those who model doxastic states using probability functions to represent partial degrees of belief. The axioms of probability theory impose constraints of probabilistic coherence on those degrees of belief, or "credences." These can be seen as a generalization of the constraints of deductive consistency and closure (jointly, "deductive cogency," as in the terminology of Christensen (2004)) that propositional logic is said to impose on our beliefs. Bayesian epistemologists may or may not impose additional normative constraints on particular credence distributions or the ways in which they should change over time and in response to new evidence. Orthodox Bayesians impose only the constraints PROBABILISM and CONDITIONALIZATION, defined in § 1.2.1 below, but I am using "Bayesian epistemology" more broadly than that to mean the sort of epistemology that traffics in probabilistic credence.

Traditional epistemology is primarily concerned with the conditions under which knowledge, or justified belief, is possible. Although it is widely accepted that Gettier (1963) shows that knowledge cannot be analysed as justified true belief, and it is also widely accepted that there is no additional, distinct[1] property of belief X such that justified true X beliefs are known, justification is still plausibly a necessary condition for knowledge. We can distinguish, as is common, between *propositional* justification and well-founded belief.

> PROPOSITIONAL JUSTIFICATION. Belief in P is justified for S at t if and only if it is epistemically permissible for S to believe that P at t.

This differs from well-founded belief in several ways; for S to have a well-founded belief in P at t, S must actually believe P at t, S's belief must somehow be based on whatever makes the attitude appropriate (in a typical evidentialist account such as that of Feldman and Conee (2004, p.83), this will mean that somehow or other S's belief must be appropriately based on or responsive to her evidence for P), and S must not be improperly ignoring counterevidence. It is propositional justification that I will take to be the fundamental notion for traditional epistemology, and which is plausibly a necessary condition on knowledge. For if S knows that P, then it is epistemically permissible for S to believe P.

It is not clear how the Bayesian and traditional approaches to epistemology fit together. One possible view is that the tools of Bayesian epistemology will allow us to give a reductive account of PROPOSITIONAL JUSTIFICATION. Call this a "probability first" approach to epistemology. A reductive account might go something like this: we describe conditions under which credence distributions are permissible, and then we analyze the doxastic attitude of belief in terms of features of the believer's credence distribution.[2] If the fact that I believe that P is reducible to facts about my credence distribution, then my being justified in believing that P will just be permissibly having a credence distribution such that I believe that P. In the following sections, I will argue against several prominent reductive accounts of justification.

Another view that Bayesian epistemologists might take is that Bayesian epistemology renders traditional epistemology pointless; it is more fruitful just to talk about credence and eschew the language of belief, disbelief, and suspension of judgment. Oft-quoted in this

[1] That is, distinct from the property of being known; obviously justified true beliefs are known if they are known, or known-if-true, etc.

[2] Weatherson (2005) is the best developed example of a reductive account. I discuss his view in §1.2.3.

regard is Jeffrey (1970, 171-172):

> By 'belief' I mean the thing that goes along with valuation in decision-making: degree-of-belief, or subjective probability, or personal probability, or grade of credence. I do not care what you call it because I can tell you what it is, and how to measure it, within limits... Nor am I disturbed by the fact that our ordinary notion of belief is only vestigially present in the notion of degree of belief. I am inclined to think Ramsey sucked the marrow out of the ordinary notion, and used it to nourish a more adequate view.

Similarly, someone who does not share Jeffrey's inclination might worry that if the concept of *knowledge* has gone missing from Bayesian epistemology, then we are no longer doing epistemology properly-so-called. Although she is not endorsing this view, Sarah Moss (unpublished talk) offers a nice articulation of the worry.

> Bayesians articulate constraints on rational credences: synchronic constraints on what credences you may have, and diachronic constraints on how your credences must evolve. Like traditional epistemologists, Bayesians are concerned with norms governing your doxastic state. But in modeling your doxastic state, Bayesians do not represent what full beliefs you have. And so they do not have the resources to talk about which of those beliefs constitute knowledge.

Although I will argue against several prominent, reductive views of the relationship between belief and credence, in § 1.2.4, I will suggest that the reason these views fail is precisely that they do not avail themselves of all the tools available to a probability first approach to epistemology. If we allow ourselves to appeal to a broader set of facts about credence distributions in attributing belief, then a reductive, probability first approach is still plausible.[3]

1.2.1 JUSTIFICATION OF BELIEF AND CREDENCE

On a "probability first" approach to epistemology, we start with a conception of justification that applies first to credences, and seek to answer questions in traditional epistemology about knowledge and justified belief using the tools of probability theory. This starts with a commitment to

> Probabilism. The doxastic state of an ideally rational subject at any time is represented by a probability function.

[3] See also Hájek and Hartmann (2010) for a more general take on the fruitfulness of building bridges between probabilistic and traditional epistemology.

Bayesian epistemologists also assume that conditionalization on new evidence is how an ideally rational subject will update her credence function over time.

> CONDITIONALIZATION. If cr_t is a subject's credence function at time t and she acquires propositional evidence E (and nothing stronger) between t and t', her credence function should become $cr_{t'} = cr_t(\cdot|E)$, where $cr_t(A|B) = \frac{cr_t(AB)}{cr_t(B)}$, if $cr_t(B) \neq 0$.

And we can speak of a "prior credence function" $cr_\emptyset$ which is not the subject's credence function at any particular time, but an idealization that represents her credences uninformed by any evidence. We can then state a condition of justification for a particular credence in a proposition.

> JC. $cr_t(P) = x$ is *justified* for a subject S at time t if and only if $\exists p$ with $p(P|E) = x$, where E is S's total evidence at time t and p is a permissible prior probability function.[4]

As I wrote above, a natural way of pursuing a probability first approach is to attempt to reduce belief or justification to believe to facts about credence. Such a reduction might be either metaphysical or epistemological. By metaphysical, I mean that whether or not a subject believes that P will supervene on her credal state—any two subjects with the same credences will have the same doxastic states. However, I am open to accounts of belief on which it is psychologically possible to impermissibly believe that P when one permissibly has credences such that one ought to believe that P. Voluntaristic accounts of belief may have this feature; even after assessing the evidence correctly, there is still a "decision" to be made about whether or not to believe that P, for example. So it may just be that whether or not it is permissible to believe that P depends on one's credal state and its epistemic status. Call this an *epistemological*, rather than metaphysical, reduction of belief to credence. The two can be made to coincide with the stipulation that, when a subject has justified credences, she believes what her credences permit her to believe. I do not claim that even ideally rational inquirers will do this in all circumstances, but for my purposes it will be harmless to impose this constraint.

Another important caveat is that I am primarily interested in the *epistemically evaluable* features of a subject, on which her doxastic states (or her permissible doxastic states)

[4] Bayesian personalists will impose no constraints on an ideally rational subject's prior probability function beyond those that are entailed by its being a probability function. Few Bayesians are this permissive; I will not be below.

supervene. As I will discuss when I consider pragmatic encroachment views in § 1.2.3, possibly whether or not one believes that P depends not just on one's credal state, but on other features of one's psychology or environment that are epistemologically neutral, such as one's preferences.[5] For example, Fantl and McGrath (2002) argues that whether or not one knows that P depends on what is at stake in acting on P; Ross and Schroeder (forthcoming) presents cases that seem to run counter to the intuitions on which the case for pragmatic encroachment depends. I find the intuitions on which both arguments depend compelling, and I do not have a settled view about whether preferences matter to (justified) belief, though I am inclined to think that they do. Thus, I will assume that any reductive account of belief includes a "holding fixed the epistemologically neutral features of the subject" clause, but (except in my discussion of the pragmatic encroachment view) I will mostly ignore this qualification.

Credences have often been called "degrees of belief," and this suggests quite naturally that at some point, degree of belief will be high enough to constitute belief. So many probability first epistemologists (and perhaps others) will want to endorse some version of the following.

> JCB. For any proposition P, a subject S with total evidence E is justified in believing that P at time t just in case $cr_t(P) = x$ is justified for S, i.e., there is a permissible prior probability distribution p such that $p(P|E) = x$ and, necessarily (if we hold fixed the epistemologically neutral features of S), if S has credence $cr_t(P) = x$, S believes that P.

It is important to be very clear about the scope of the operators in JCB. So, let $\triangledown$ indicate epistemic permissibility, and let $\Box$ indicate necessity (whatever kind of necessity is involved in the supposed reduction of belief to credence). Let $Bel(P)$ indicate that S believes that P at t. The intended reading of JCB is:

$$\forall P(\triangledown Bel(P) \equiv \exists x(\triangledown cr_t(P) = x \;\&\; \Box(cr_t(P) = x \supset Bel(P)))).$$

This is very different from the following.

$$\forall P(\triangledown Bel(P) \equiv \exists x(\triangledown cr_t(P) = x \;\&\; \Box\forall Q(cr_t(Q) = x \supset Bel(Q)))).$$

[5] This is not a perfectly precise characterization. I will suggest in § 3.3.2 that in fact ideal epistemic rationality requires a desire to understand, or curiosity about unexpected phenomena. But I am not sure how to make it more precise at this juncture.

This second principle is false, as I shall show in the next section, in the course of arguing against a pair of reductive views that entail it. JCB is more plausible. However, the most prominent accounts that endorse JCB are *pragmatic encroachment* views on which the reason it is possible to have $cr_t(P) = cr_t(Q)$ and yet be justified in believing P but not Q is that the practical stakes that hinge on P matter to the level of credence on must have in P in order to count as believing it.[6] In §1.2.3, I argue that pragmatic considerations do not suffice to explain every such case. In particular, they do not suffice to explain the case that I will present as a counterexample to the THRESHOLD view in the next section.

1.2.2 THE THRESHOLD VIEW

A simple, natural way of spelling out when credence suffices for belief would simply appeal to a threshold.

> THRESHOLD. There is some $\beta \in (.5, 1]$ such that, for any subject S, time t, and proposition P, S believes that P at t just in case $cr_t(P) \geq \beta$.

Presumably β will be fairly high, but less than 1. For one may certainly believe a proposition of which one is uncertain (if we treat maximal subjective probability as certainty), or on which one would not bet one's life in order to win a dollar, if we adopt the usual story about how a subject's credences function in her decision-making. If we are going to expect rational beliefs to be at all consistent, we will also want β to be greater than .5 (probably much greater), or the THRESHOLD view will entail that we sometimes believe pairs of contradictory propositions.

The THRESHOLD view has been explicitly defended by Weintraub (2001), and more recently by Sturgeon (2008).[7] Unfortunately, this view requires us to abandon deductive

[6] Cf. Fantl and McGrath (2002, 2007); Hawthorne (2004). Only Weatherson (2005) is explicit in separating the reductive analysis into an epistemic component (the norms on permissible credence) and a non-epistemic component (the conditions under which a particular subjective credence, justified or not, entails belief). Not uncoincidentally, he is also the only writer of the bunch who explicitly adopts a probabilistic approach. But I think that all of the other views are compatible with JCB, and this sort of reductive analysis of belief is a natural way of spelling it out.

[7] Weintraub (2001) actually defends a slightly different proposition that directly connects "rational acceptability" of propositions with rationally ascribed credence. I think it is fair to attribute to her JCB and THRESHOLD. "There is a threshold, $\epsilon < 1$, such that a proposition is rationally acceptable if its subjective probability is greater than ϵ and is itself rationally ascribed" (Weintraub, 2001, 439).

cogency for rational belief. This is, needless, to say, a pretty radical move. But it is required by THRESHOLD because it is quite possible to find, for any value of β, an x slightly higher than β such that $x^2 < \beta$. But if P and Q are probabilistically independent, then $cr(PQ) = cr(P)cr(Q)$. So if $cr(P) = cr(Q) = x$, and the threshold for belief is β, then P and Q will be believed but PQ will not be, in violation of closure.

However, I find THRESHOLD implausible even if the preface paradox convinces us to deny multi-premise closure for binary belief.[8] It is enough if the intuitive verdict about the following case is correct.

> CARD CASE. Suppose that I show you a deck of n cards, and show you that it contains exactly one Ace. I shuffle the cards thoroughly, and you cut the deck. You know I am doing this in order to construct a real life case with which to make a philosophical point about credence, and we have no practical stake in which card is drawn, so you know I have no reason to cheat or perform any feats of card trickery (not that I even know how to). I am about to draw a card and show it to you.

Let A be the proposition that the Ace will be drawn. Then:

- You should have $cr(A) = \frac{1}{n}$.
- You should neither believe nor disbelieve A (i.e., you should suspend judgment in A).

This is quite obvious for $n = 2$, and other small values of n. Now, if you should have credence $\frac{1}{n}$ in A, your credence in $\neg A$ should be $1 - \frac{1}{n}$. But, for large enough n, $1 - \frac{1}{n} > \beta$ for any $\beta \in (.5, 1)$. So the THRESHOLD view entails that whether or not you should believe $\neg A$ depends on the size of the deck of cards. But $\neg A$ is a paradigmatic lottery proposition—i.e., a proposition that is a member of a set of exhaustive and mutually exclusive possibilities such that we know that one will or does obtain and such that we have no reason to treat any one possibility as more or less likely than the rest. The intuition that we do not know that any particular possibility will not obtain, no matter how probable, is robust in this kind of case even without exploiting closure to draw a *reductio*.[9] Indeed, the point of the lottery paradox is that we can exploit the strength of this intuition to argue that *other* propositions that are intuitively known are not, in fact.

[8] See Christensen (2004); Hawthorne (2004) for thorough discussion of these paradoxes and their consequences for closure.

[9] This sort of case also motivates the Principle of Indifference in probabilistic epistemology, which says that when we have a partition of cases of this kind, we should give them equal credence.

Justification to believe is a less intuitive notion than knowledge, and so, possibly, the intuitive judgment that you do not *know* $\neg A$ will not carry over to the claim that you are not justified in believing $\neg A$. Nevertheless, it is clear that it is *both* obligatory to have $cr(\neg A) = 1 - \frac{1}{n}$ in COIN CASE *and* at least permissible to suspend judgment in $\neg A$. This shows that justified belief is not reducible to justified credence above some *fixed* threshold, since $\exists n, \forall x \in (.5, 1)$, such that $1 - \frac{1}{n} > x$.[10]

Defenders of THRESHOLD can respond to this by positing contextual variability. That is, they can say that in any context, THRESHOLD is true. However, between contexts, the threshold may move up or down. The view that the truth conditions for "S knows that P" vary depending on the interests and/or the possibilities salient to the individual ascribing knowledge to S has been an influential view in traditional epistemology quite apart from any concern for the relationship between credence and belief. The contextualist views about knowledge in DeRose (1992, 1995) and Cohen (1998) allow standards to vary only with respect to the context of ascription.

> We can think of the context-sensitivity of knowledge ascriptions in this way: For each context of ascription, there is a standard for how strong one's epistemic position with respect to proposition P must be in order for one to know P. Where two contexts differ with respect to this standard, a speaker in one context may truly say "S knows P," while a speaker in the other context truly says "S does not know P." (Cohen, 1998, 289)

I do not want to assume that the defenders of such contextualism about knowledge ascriptions would accept a probability first, reductive approach to the relationship between credence and belief. But it is clearly compatible with such a view. On this sort of approach to contextualism, THRESHOLD is true in every context, but in some very demanding contexts, perhaps, $\beta = 1$. In fact, the contextualist may well say that when we think about lottery-style cases like CARD CASE, we shift the standards for belief-ascription so that we will only ascribe belief in near-certainties, where "near" can be as close as we like (by thinking of bigger and bigger lotteries). Thus, we can explain the intuitive judgments I offer about

[10] I am assuming a non-permissive epistemology in which it is never the case that both belief that $\neg A$ and suspension of judgment are permissible at the same time, for a single subject. I think in fact that one should not believe $\neg A$ in CARD CASE. But I do not need such a strong claim to get the conclusion—even if the permissive view is true, I can show that THRESHOLD is false with only the assumption that there is just one permissible credence. I will in fact endorse a limited permissivism in **??**, but this permissivism holds at the level of credence, not belief: sometimes the evidence does not suffice to determine a unique permissible credence. CARD CASE is not such a permissive case, however.

CARD CASE, while still maintaining that in most ordinary contexts, the threshold for belief is less than 1.

My argument against this view will proceed by showing that it is possible to combine CARD CASE with a COIN CASE in such a way that, within the same context, S should have (where H is the proposition that the next flip of the coin will be heads) $cr(H) = cr(\neg A)$ and yet she should both believe H and suspend judgment about $\neg A$. Here's the case I have in mind.

> COIN CASE. Suppose that I show you a coin. I tell you that this coin was selected at random from a bag of coins in which $\frac{1}{3}$ of the coins were fair, $\frac{1}{3}$ were completely biased towards heads (they always land heads), and $\frac{1}{3}$ were completely biased towards tails. I do not know which kind of coin it is. I will flip the coin several times, in order for us to get some evidence about which kind of coin it is.

Let H be the proposition that the next flip of the coin lands heads. Let B be the proposition that the coin is completely biased towards heads, and let F be the proposition that the coin is fair. Let cr_n be your credence function after n flips. Your conditional credences in H given each hypothesis about bias should not change no matter what evidence you get. So for all n:

- $cr_n(H|B) = 1$;
- $cr_n(H|F) = \frac{1}{2}$;
- $cr_n(H|\neg(B \vee F)) = 0$.

Your credence in H after n trials should be the average of the conditional credences above, weighted according your credence in each bias hypothesis. That is:

- $cr_n(H) = cr_n(B) + \frac{1}{2} \cdot cr_n(F)$.

Your conditional credences in each bias hypothesis given H should be given by Bayes' Theorem, which yields:

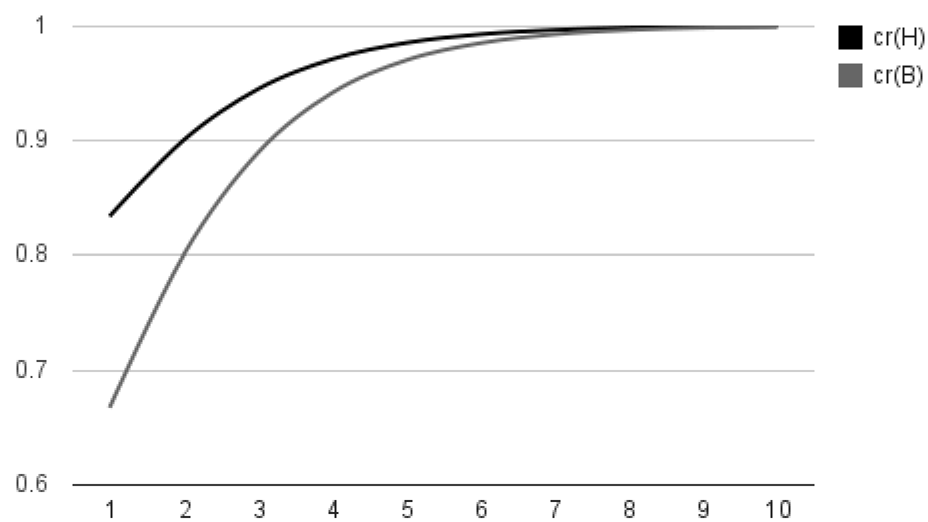
- $cr_n(B|H) = \frac{cr_n(B)}{cr_n(H)}$;
- $cr_n(F|H) = \frac{1}{2} \cdot \frac{cr_n(F)}{cr_n(H)}$;
- $cr_n(\neg(B \vee F)|H) = 0$.

Suppose that, as I flip the coin, we get nothing but heads results. Given an assumption of equal initial credence of $\frac{1}{3}$ in each bias hypothesis, if you update via CONDITIONALIZATION and get no other relevant information about the coin,[11] your credences will be given by the following formulae, for $n > 0$.

- $cr_n(B) = \frac{2^{n+1}}{2+2^{n+1}}$;
- $cr_n(H) = \frac{1+2^{n+1}}{2+2^{n+1}}$.

We can see (Figure 1) that B will quickly be confirmed, and of course you will always be more confident that H than B, since B entails H. Although $cr_n(H)$ and $cr_n(B)$ asymptotically approach 1, neither will ever equal 1 (and of course, after any sequence of consecutive heads results, a single tails result will cause $cr_n(H)$ to collapse to $\frac{1}{2}$ forevermore, since F will then be confirmed with certainty).

Figure 1: CREDENCE AFTER N CONSECUTIVE HEADS TRIALS



Now, H and B are not lottery propositions. I claim that it is intuitively clear that, for some n, a sequence of n consecutive heads results constitutes sufficient evidence for you to

[11]As it happens, I do not agree with the Bayesian orthodoxy that we should always conditionalize, for reasons such as those given in Arntzenius (2003) and Weisberg (2009). However, this is as clear a case as there is in which we absolutely *should* update by conditionalization.

be justified in believing both B and H. After even 10 consecutive heads results, for example, you ought to regard a bet that paid \$1 if H and cost \$2,000 if $\neg H$ as a good bet (it would be evaluated as fair for $cr_{10}(H) = \frac{1+2^{11}}{2+2^{11}}$ if it cost \$2,050). At that point you should be very nearly certain that the next flip will come up heads. If we were looking for a fair coin from the bag, we'd want at that point to put the coin down and draw another one to test. It would not be worth your time to keep flipping the coin for further evidence, unless something *very very important* was riding on knowing whether this particular coin was biased—but let's hypothesize that in COIN CASE there are no such stakes.

This is consistent with the contextualist view. For, no matter what the contextually-determined threshold β is, there will be some number n such that n consecutive heads flips would suffice to justify belief in H. But now we can combine the two cases presented in this essay to get a counterexample. For suppose that, after n trials, we stop flipping the coin and I pull out a deck of $2 + 2^{n+1}$ cards. Then, if the intuitive judgments defended above about CARD CASE are correct, you should have precisely $cr_n(\neg A) = \frac{1+2^{n+1}}{2+2^{n+1}}$, and you should suspend judgment. But this is the exact same credence that suffices for belief in H. I claim that nothing in this story should cause us to think that the threshold has somehow shifted so that either our intuitive judgment about CARD CASE should change or so that, after I pull out the deck of cards, you should cease to believe H. If so, then within a single context, you should have $cr(\neg A) = cr(H)$ and suspend judgment about $\neg A$ while believing H. This shows that the contextualist defense of THRESHOLD fails.

The soundness of this argument obviously depends on the claim that there is no contextual variation between COIN CASE and CARD CASE. It seems obvious to me that there *need* not be any relevant difference in the standards by which we evaluate these beliefs. But nevertheless, let me try to defend this claim.

Hawthorne (2004, 55–56) offers the following as a natural way of extending a contextualist semantics for such terms as "flat" to "know."

> Whether a true belief that p counts as knowledge depends, it would seem, on how easily that believer could have been mistaken about p and about similar subject matter. But how easily is easily enough? Which possible mistakes that a person could have made about p and about similar subject matter defeat a knowledge ascription? In general, the easier it is for a subject to make a particular mistake, the more semantic pressure there is to count that mistake relevant to the knowledge claim. But what determines the threshold? Let our

> toy semantic theory say that an ascription of the form "S knows that p" is true on some occasion iff there is no possible world in which S makes a mistake that is relevant on that occasion. And let the standards of relevance vary from ascriber to ascriber.

While the story Hawthorne tells here is about when a mistake is "relevant" to whether a true belief is known, rather than justified, I think it is plausible that the contextualist take on THRESHOLD could be defended in the same way. If the question is, "in which contexts is it permissible to believe that P, holding fixed the value of $cr(P)$?" one way to explain the difference would be by appealing to a difference in the relevant mistakes one might make in believing that P. If the danger of making a mistake is relevant, then one should not believe that P, whereas if it is irrelevant, one may believe that P. One the reductive story, this difference must wind up being a difference in when one *does* believe that P, rather than when one *should*. My real interest, however, is in how the credences one ought to have relate to the beliefs that one ought to have. Ultimately, it is of no matter to this epistemological story whether it is psychologically possible to believe that P when one epistemically ought not, given one's (permissible) credences, as it would perhaps on some voluntaristic accounts of belief.

The way I have told the story, you and I (both as characters in the story and as ourselves) are the only ascribers to be concerned with. It is possible, for all I know, that you are currently really worried about the possibility of being deceived by an evil demon, or some other skeptical hypothesis that would undermine your justification for H in COIN CASE, but let me just stipulate that things are normal, so we do not have to worry about that. The way I have constructed the case, the only ways for you to have a mistaken belief in H are for the coin to be biased towards tails or fair. The former is certainly false, given your evidence in the case, so it is not a relevant mistake. The latter is not ruled out with certainty by your evidence. But if you agreed with my intuitive judgment that, after some number n of consecutive heads results, you are justified in believing that H, then you've agreed not to count the possibility that $\neg HF$[12] as a relevant mistake in the context. I am not exactly sure how to argue that your standards for evaluating whether belief in H is justified *did not* shift after I asked you to consider the possibility that I pull out a deck of $2 + 2^{n+1}$ cards. But it

[12] The negation has narrow scope here; if I meant the negation of the conjunction I would write $\neg(HF)$.

seems to me that they need not have.

Of course, I do not deny that you should think that you are *not* justified in believing that $\neg A$ in CARD CASE. If you are a committed contextualist, what you'll say is that *because* you treat the possibility that A as a relevant possible mistake in CARD CASE, when you are considering CARD CASE you have to treat any other proposition that you take to be just as risky to be a relevant possibility as well. Since $\neg HF$ has precisely the same subjective probability for you, you have to treat it as relevant.[13] It is only fair, or something.

Against this, consider that I could easily include a second Ace in the deck of cards. Suppose I did that. Then your credence in $\neg A$ should be $\frac{2^{n+1}}{2+2^{n+1}}$. Note that this is precisely what your credence in B should be after n trials, so you will have $cr_n(\neg A) = cr_n(\neg F)$. If I used two aces instead, then, H may still meet the threshold for justified belief, but B will not. Using the contextualist explanation for why not, we will have to say that F is risky enough to count as a relevant possibility, but $\neg HF$ is not.

I claim that this is implausible. For if it is a relevant possibility that the coin is fair, then surely it is not plausible that we can ignore as irrelevant the possibility that it comes up tails. The notion of "relevance" invoked by the contextualist is unfortunately vague, but we can precisify things a bit by considering a couple of possible interpretations. In the quotation above from Hawthorne (2004), he talks of relevant mistakes as relevant possible worlds, which suggests that relevance will be cashed out as an accessibility or closeness relation. But if HF is accessible or close enough, it is difficult to see a non-*ad hoc* way of defending the view that $\neg HF$ is not also accessible or just as close. Alternatively, I have suggested that the contextualist might explain relevance in terms of subjective risk assessment. But if F is sufficiently risky, and, as is clear, $cr_n(H|F) = cr_n(\neg H|F)$, again it is hard to see how to motivate the claim that there is a relevant difference between $\neg HF$ and HF under the circumstances. If the contextualist cannot provide an explanation of this difference, then her last alternative is to claim that introducing the deck with two Aces still raises the standards to the same degree that the deck with one Ace did. But why on earth would it do that?[14]

[13] Note that the way I have constructed the case, you the reader should have precisely the same credences conditional on being in that case as the you the character in the case should have.

[14] Ok, I can think of two ways. First, it could happen by introducing a relevant possibility about *another* card of which there is only one in the deck. So stipulate that there are at least two of each card in the deck. Second, as I said above, the contextualist might claim that introducing a lottery proposition does not just

The contextualist has a possible response here, in the form of a two-stage story. She might say that first, when I produce the deck of cards with two Aces, the threshold for belief is raised above $\frac{2^{n+1}}{2+2^{n+1}}$, but possibly not higher than $\frac{1+2^{n+1}}{2+2^{n+1}}$. At this point, belief in B, and hence in $\neg F$, is impermissible, because F is risky enough to count as a relevant mistake, but $\neg HF$ may not yet be risky enough to render belief in H impermissible. However, there is a second stage, in which we notice that F and $\neg HF$ are relevantly similar, and that the same doxastic attitudes must be permissible with respect to both. Since belief in F is not permitted in the context, this forces the standards for belief up yet again, so that $\neg HF$ counts as a relevant mistake as well, and belief in H is also not permitted. That is, in any context including COIN CASE, it is not possible for the threshold to be within the interval $[\frac{2^{n+1}}{2+2^{n+1}}, \frac{1+2^{n+1}}{2+2^{n+1}}]$. And in any context including CARD CASE (with two Aces), the threshold must be at least $\frac{2^{n+1}}{2+2^{n+1}}$. These restrictions combine to force the threshold to at least the point that it would be in CARD CASE with a single Ace.

I doubt I can convince a committed contextualist defender of THRESHOLD not to embrace this line of defense. But I do not see any motivation for thinking that these restrictions on the threshold for H and B and the threshold for $\neg A$ should *combine* in this way, other than a commitment to THRESHOLD. Even if we agree that the same doxastic attitude to H and B should be permissible in COIN CASE and that belief in $\neg A$ is impermissible in CARD CASE, it does not immediately follow that belief in H and B is impermissible in a context that includes both. Moreover, I think there are significant differences between the cases that make belief in H and B intuitively more appealing than belief in $\neg A$, even after considering CARD CASE. I will elaborate in § 1.2.4, but the basic idea is that in COIN CASE, we the evidence that favors B in fact also supports belief that we are safe from making a mistake if we believe H. There is no similar support, in CARD CASE, for the belief that we are safe from making a mistake if we believe $\neg A$. So we should not accept even a contextualist form of THRESHOLD.

raise the threshold to $1-x$, where x is the smallest relevant chance of winning the lottery. Instead, it raises it all the way to 1. But it certainly seems to me that if the deck of cards contains just a few cards—say, less than 5—there is *no* temptation to say that the context has shifted so that we believe only certainties. The size of the lottery matters, and the most plausible way for it to matter is to raise the threshold to $1 - \frac{1}{n}$ when there are n possible outcomes.

Now, Hawthorne (2004) defends the possibility of letting the standards for knowledge vary not with respect to the possible mistakes that are salient to the ascriber of knowledge, but with the possibilities salient to the subject of the ascription—he calls this view "sensitive moderate invariantism." I have two things to say about this view. The first is short, the second will take up the rest of this essay. Hawthorne initially describes this character as follows.

> I earlier introduced the sensitive moderate invariantist as one who claimed that the extension of "know" "depends not merely upon the kinds of factors traditionally adverted to in accounts of knowledge—whether the subject believes the proposition, whether that proposition is true, whether the subject has good evidence, whether the subject is using a reliable method, and so on—but also upon the kinds of factors that in the contextualist's hands make for ascriber-dependence." (Hawthorne, 2004, 173)

Hawthorne attempts to explain what these extra factors are in terms of the concept of salience of possibilities—a possible mistake becomes relevant by being made salient to the ascriber, or the subject for an sensitive moderate invariantist. Insofar as the *only* innovation the sensitive moderate invariantist makes is to allow the variation in standards to be sensitive to features of the subject, I expect that the argument above will apply just as well to the sensitive moderate invariant interpretation of THRESHOLD.

> SENSITIVE THRESHOLD. For each subject S and time t, there is a $\beta \in (.5, 1]$ such that, for all propositions P, S believes that P at t just in case $cr_t(P) \geq \beta$.

If you have been convinced by my argument that you should ascribe to yourself belief in H but not in $\neg A$ despite giving them equal credence, in COIN CASE and CARD CASE, then you are going to be hard pressed to convince me that there is a possibility that is salient to the character you in the story I told that is not salient to you, the reader.

However, there are sensitive moderate invariantists who have a more sophisticated story to tell about how the standards for justified belief vary with respect to the subject of the belief ascription (and in fact the limited sensitive moderate invariantism described above may have no actual defenders). Hawthorne (2004, §4.4) argues that an appeal to salient possibilities is limited, and insufficient to the epistemological task insofar as it is sometimes the case that possibilities that *should be* salient to someone are not. He then proposes that the subject's practical interests should be taken to play a role in determining which possibilities are relevant.

> We now have the outlines of a second mechanism that may be introduced by the sensitive moderate invariantist. The basic idea is clear enough. Insofar as it is unacceptable—and not merely because the content of the belief is irrelevant to the issues at hand—to use a belief that p as a premise in practical reasoning on a certain occasion, the belief is not a piece of knowledge at that time. (Hawthorne, 2004, 176)

A similar view has been defended by Fantl and McGrath (2002, 2007). However, this view seems to me to already commit us to the denial of the SENSITIVE THRESHOLD view. For it is clear that it may well be appropriate to use H in practical reasoning though it is not appropriate to use $\neg A$ in practical reasoning, even within the same context. Consider the following bits of reasoning.

> We have flipped the coin 10 times, and I have shown you a deck of 2050 cards. You purchase fair bet on A, which pays \$2,050 if the Ace is drawn and cost you only \$1. I now offer to buy the bet from you for a penny. You reason as follows.
>
> 1. The Ace will not be drawn.
> 2. If I sell my bet and the Ace is not drawn, I get a penny.
> 3. If I do not sell my bet and the Ace is not drawn, I get nothing.
> 4. So, I will sell my bet.

This is a terrible piece of practical reasoning. But nothing of similar interest is riding on H or B. In fact, consider the following offer.

> We've flipped the coin 10 times, and I have shown you a deck of 2050 cards. You purchase fair bet on A, which pays \$2,050 if the Ace is drawn and cost you only \$1. I also offer to let you borrow the coin. You reason as follows.
>
> 1. The coin is biased towards heads.
> 2. If the coin is biased towards heads, I can make money by telling my gullible sister that it is fair and offering her a bet that pays \$5 if tails and costs \$1.
> 3. I should borrow the coin.

This is *not* a terrible piece of practical reasoning, even after you've purchased the bet on A. So this pragmatic version of sensitive moderate invariantism already conflicts with the THRESHOLD view. In § 1.2.3, I will argue further that, while there may be pragmatic encroachment of this kind on epistemic standards, it cannot be the whole story—this sort of pragmatic variation cannot account for every case in which $cr(P) = cr(Q)$ but one is justified in believing P but not Q.

1.2.3 PRAGMATIC ENCROACHMENT

The THRESHOLD view entailed the second, unintended interpretation of JCB in §1.2.1, which I formalized as follows.

$$\forall P(\triangledown Bel(P) \equiv \exists x(\triangledown cr_t(P) = x \ \& \ \Box\forall Q(cr_t(Q) = x \supset Bel(Q)))).$$

On this sort of interpretation, when a particular possible credence x suffices for belief in P, it also suffices for belief in any other proposition Q. I have argued that the pragmatic version of sensitive moderate invariantism must deny this. But such sensitive moderate invariantism is still consistent with JCB, although among its defenders, only Weatherson (2005) explicitly adopts a reductive, probability first approach.

Fantl and McGrath (2002, 77) defends the following pragmatic condition on justification to believe.

> PC. S is justified in believing that P only if S is rational to prefer as if P.

The attitude of "rationally preferring as if P" is explained this way:

> For any states of affairs A and B, S is rational to prefer A to B, given P, iff S is rational to prefer A to B, in fact.

This principle explains intuitive differences between cases in which it seems that S knows that P because nothing much rides on whether P is true and cases in which S' does not know that P because something of importance to her rides on it (so that, e.g., S would not reasonably continue to collect evidence about P but S' ought to). I appealed to this intuitive difference in COIN CASE above when I noted that, after 10 consecutive heads, we'd be rational to stop flipping the coin and take it as settled that it is biased *unless* something of extreme importance hinged on knowing whether that particular coin was fair or biased. But the principle does this without sacrificing the intuitive distinction between epistemic justification for believing P and pragmatic justification for believing P, for it may well be that one is rational to prefer as if P when the costs of believing P are nevertheless high (such cases are easy to construct simply by positing a coercive threat of severe injury if one fails to believe that P).

I claim that PC cannot explain the difference in what you should believe in COIN CASE and CARD CASE. *If* something is riding on whether the Ace is drawn but nothing much is riding on whether the next flip of the coin is heads, as in the examples of practical reasoning at the end of the last section, then PC gives the right result: you are justified in believing H but not $\neg A$, because you have no preferences that would change conditional on H, but you do have preferences that would change conditional on $\neg A$ (in my sample reasoning, you prefer selling your bet for a penny conditional on $\neg A$). But in the original presentation of the cases, nothing turned on either proposition. I claim that it is still intuitively the case that you should suspend judgment that $\neg A$ and believe H. But there is no preference that you have, in CARD CASE considered without additional trappings, that changes if we assume $\neg A$.

Of course, you can point to *possible* preferences that you might have that are not invariant given $\neg A$. For example, if B is accepting a bet that pays \$2,051 if the Ace is drawn and costs \$1, then you may rationally prefer B to $\neg B$ if the deck contains 10 cards, but given $\neg A$, you do not. So you do not rationally prefer as if $\neg A$ when that bet is offered. This is why the pragmatic encroachment view allows for contextual variation; which possible outcomes matter to the question whether it is rational to prefer as if P depend on which are live options. Thus, my claim is that, intuitively, if there is nothing riding on H or $\neg A$ for you other than satisfying your curiosity or deciding what to believe given the evidence, as there may easily not be, you should believe H after, say, 10 consecutive heads flips while suspending judgment about $\neg A$.

That is to say, I think H and $\neg A$ can, in the right context, constitute a pair of *pragmatically irrelevant* propositions to which you should give equal credence but adopt different doxastic attitudes. Fantl and McGrath (2002) does not explicitly address what to say about pragmatically irrelevant propositions, but Weatherson (2005) does. Weatherson spells out the pragmatic condition on belief ascription in a similar way.

> An agent believes that P iff conditionalising on P does not change any conditional preferences over things that matter.

In symbols: $Bel(P) \equiv \forall A \forall B \forall Q\ (A \succeq_Q B \equiv A \succeq_{PQ} B)$, where $A \succeq_Q B$ indicates the relation of preferring A to B conditionally on Q. That is to say, S believes that P just in case adding P to her evidence changes none of her conditional preferences, so that assuming P would not make a difference to her actions. We can substitute this definition into JCB.

$$\forall P(\triangledown Bel(P) \equiv \exists x(\triangledown cr_t(P) = x \;\&\; \Box(cr_t(P) = x \supset \forall A \forall B \forall Q\ (A \succeq_Q B \equiv A \succeq_{PQ} B)))).$$

This says that it is permissible for S to believe that P just in case it is permissible for her to have credences such that $cr_t(P) = x$ and necessarily, given that credence (and holding fixed the domain of the quantifiers, i.e., holding fixed the "things that matter" to S), S's conditional preferences are the same as they would be conditional on P as well. This will be true in just those circumstances in which your credence is either very high, or the things that both matter and depend on P do not matter very much. For example, if the only live option is whether to purchase a bet on P that costs \$2 and pays \$1 if P is true, then as long as your credence in P is greater than $\frac{2}{3}$, you will prefer to take that bet, and you would prefer to take that bet conditionally on P, so you believe that P. But if the stakes were higher, you would need to have a correspondingly greater credence in P to count as believing that P.

Weatherson is also more perspicuous about what the "things that matter" are.

> An action A is a live option for the agent if it is really possible for the agent to perform A. An action A is a salient option if it is an option the agent takes seriously in deliberation.. . . We'll say the two initial quantifiers range over options that are live and salient options for the agent.

Given this definition of belief, Weatheron's view entails that we believe *both* P and $\neg P$ for any practically irrelevant P, such as the proposition that Julius Caesar ate breakfast before crossing the Rubicon. This proposition makes no difference to any choice that we might have to make—it would even be hard to construct an interesting bet on this proposition to make it practically relevant, since there is presumably no way of deciding who wins the bet.[15] To

[15] Is this a relevant difference between my propositions H and $\neg A$ and Weatherson's example? Does the *possibility* of constructing decidable bets on H and $\neg A$ somehow affect whether we should prefer as if they're true? No. For, first, if it did, it would have to affect both propositions in the same way, but again, it seems quite intuitive that you should believe H in COIN CASE but suspend judgment about $\neg A$ in CARD CASE. Second, we could easily change our already hypothetical cases to make them even more practically irrelevant by stipulating that they happened in the past. Suppose, for example, that we flipped the coin 10 times and then stopped, and shuffled the deck of 10 cards, but after considering H and $\neg A$ and before we could flip/draw, the coin and the deck were then locked away in a sealed vault and deposited at the bottom of the Marianas trench. We can still consider what credence we should have had in H and $\neg A$ and what the appropriate doxastic attitudes were.

solve this problem, Weatherson modifies his view to include, among the things that matter with respect to P, the "actions" of believing and not believing P and other relevant, salient propositions. [16] He then stipulates that we prefer to believe that P when we have credence in P greater than $\frac{1}{2}$. He further stipulates that for any relevant or salient propositions P and Q, believing/not believing P and believing/not believing Q are open and salient "actions" with respect to P. This definition solves the problem Weatherson was concerned with, since it requires that you have credence greater than $\frac{1}{2}$ in P if you believe that P. Otherwise, you will not prefer to believe that P, but conditionally on P you will prefer to believe P, so your preferences are not invariant conditionally on P. So it is not possible to believe both P and $\neg P$, for any P.

However, this definition give the wrong result in COIN CASE. For H, B, and F are all clearly salient propositions. Suppose we have had a streak of 10 consecutive heads tosses. At that point, you strongly prefer to believe H, but it is not yet clear that you *do* believe H by Weatherson's defintion. Let J be the action of believing H and K be the action of not believing H. Conditional on F, your credence in H is $\frac{1}{2}$, so conditional on F, you do *not* prefer J to K. However, conditional on HF, of course, your credence in H is 1, so, conditional on HF, you prefer to believe H. But J and K are open and salient with respect to H in the stipulated sense. So, this definition entails that you do *not* believe H in COIN CASE—regardless of whether anything is riding on H except the satisfaction of your curiosity about the case, or deciding what to believe given the evidence.

Weatherson's stipulations have some attractive features that I will not go into here, but they may seem unnecessarily roundabout from the perspective of handling pragmatically

[16] The final formal definition of belief is as follows:

- A proposition P is *eligible for belief* if it satisfies $\forall A \forall B \forall Q$ $(A \succeq_Q B \equiv A \succeq_{PQ} B)$, where the first two quantifiers range over the open, salient actions.
- For any proposition P, and any proposition Q that is relevant or salient, among the actions that are (by stipulation!) open and salient with respect to P are *believing that* P, *believing that* Q, *not believing that* P, and *not believing that* Q.
- For any proposition, the subject prefers believing it to not believing it iff (a) it is eligible for belief and (b) the agent's degree of belief in the proposition is greater than $\frac{1}{2}$.
- The previous stipulation holds both unconditionally and conditional on P, for any P.
- The agent believes that P iff $\forall A \forall B \forall Q$ $(A \succeq_Q B \equiv A \succeq_{PQ} B)$, where the first two quantifiers range over all actions that are either open and salient *tout court* or open and salient with respect to P (as described above).

irrelevant propositions. It would suffice to stipulate that, $\forall P$, one believes that P only if one's credence in P is greater than $\frac{1}{2}$. Such a stipulation would have some counterintuitive consequences as well, though. Suppose, for example, that I have a deck of 1001 cards, of which 501 are Aces and the rest are not. Then I should have credence of about .5004995 in the proposition A that an Ace will be drawn. If that proposition is pragmatically irrelevant, then this modified proposal would entail that I believe it, and that I ought to believe it. I have an extremely strong intuitive reaction that this is incorrect. Adding a single non-Ace to the deck should not, for example, change the appropriate doxastic attitude.

Thus, the best developed accounts of pragmatic variation in when justified credence suffices for justified belief do not deliver the correct results about COIN CASE and CARD CASE.

Possibly, one might try to combine the threshold and preference invariance views, and say that one believes that P just in case conditionalizing on P does not change one's conditional preferences over things that matter *and* one's credence in P is above a contextually-determined threshold. However, given that THRESHOLD does not do justice to the intuitive verdicts in COIN CASE and CARD CASE, and pragmatic factors are unable to explain the difference, it does not seem promising to suppose that some combination of the two will do the trick. In the next section, I will make a suggestion about what I think goes wrong with this sort of view, and point to some other features of credence distributions which might help us explain our different reactions to COIN CASE and CARD CASE, and therefore rescue a probability first view.

1.2.4 PROBABILITY FIRST?

Given the failure of THRESHOLD and the attempt to reduce belief to invariance in conditional preference, it might seem that the right response is to abandon JCB and the probability first approach to epistemology altogether.[17] But that conclusion is premature.

For one thing, the pragmatic view that reduces belief to invariance in conditional preference is not necessarily the only way of spelling out JCB. Once we grant that, possibly

[17] Ross and Schroeder (forthcoming) offers a non-reductive account of belief that they claim also explains the cases that motivate the sensitive moderate invariantists, for example.

$cr_t(P) = cr_t(Q)$ and the former suffices for belief but the latter does not, we can accommodate my CARD CASE and COIN CASE. But JCB, on the intended interpretation, allows for that. This is unsatisfying insofar as there is no corresponding explanation of *why* credence of x might suffice for belief in one proposition and not in another.

However, I am not optimistic that we can preserve JCB simply by substituting a better account of "S believes that P." Nevertheless, we can modify JCB without abandoning a reductive probability first epistemology. While I do not have a full account to offer of when a subject's credences entail that she believes that P, I do think we can go some way towards diagnosing why the THRESHOLD and preference invariance accounts of this relationship fail, in a way that will reveal that a probability first epistemology might do better.

The main point to be observed is that JCB accounts of the relationship between credence and belief expect there to be a necessary connection between the fact that $cr_t(P) = x$ and the fact that S believes that P (holding fixed the epistemologically neutral features of S). But this is to collapse all of the information that S's credal state contains into a single value, necessarily discarding much of it in the process. This is a mistake, as many sophisticated Bayesians have been aware.

> At an intuitive level, the total evidence for a proposition X is the sum of all those considerations that tell in favor if its truth. Bayesians, and their opponents, have often proceeded as if the total *amount* of evidence for X is directly reflected in X's credence. When $c(X) = x \ldots$, this amounts to the claim that the number x is a meaningful measure of the total amount of evidence for X. More generally, the view is that (a) the person has more evidence for X than for Y iff $c(X) > c(Y) \ldots$, (2) she has strong evidence for X iff $c(X) \approx 1 \ldots$, (c) E provides the person with incremental evidence for X iff $c(X|E) > c(X) \ldots$, and so on. This picture of the relationship between credencese and evidence is seriously misleading. (Joyce, 2005, 158)

Joyce goes on to distinguish several features of the evidential support a proposition might enjoy—balance, weight, and specificity. He identifies the "balance" of the evidence with "how decisively the data tells for or against the hypothesis" (Joyce, 2005, 158), i.e., with the value of $cr_t(P)$. So in Joyce's terminology, we can complain that JCB assumes that whether or not S believes that P is determined just by the balance of S's evidence for P (and maybe her conditional preferences). This neglects other features of her credal state that reflect differences of evidential support—weightier or more specific evidence for P.

So, possibly, whether or not S believes that P does not supervene on the fact that $cr_t(P) = x$ (and S's conditional preferences), but whether or not S believes that P does supervene on the fact that S's credences are represented by cr_t (and S's conditional preferences). So we might modify JCB to allow us to appeal to more of this information.

> JCB*. For any proposition P, a subject S with total evidence E at t is justified in believing that P at time t just in case there is a permissible prior probability function p such that $p(\cdot|E)$, and, necessarily, if S has credences $cr_t = p(\cdot|E)$, S believes that P.

The fact that S believes that P and S does not believe that Q might be reflected in features of S's credal state other than the fact the values $cr_t(P)$ and $cr_t(Q)$. So a probability first approach to epistemology can work if JCB* is true, even if JCB is not.

So what is the "weight" of the evidence, and how does it allow us to explain the difference between S's justified belief in H in COIN CASE and permissible suspension of judgment about $\neg A$ in CARD CASE?[18]

> As the relevant evidence [for a hypothesis] at our disposal increases, the magnitude of [its] probability may either decrease or increase, according as the new knowledge strengthens the unfavorable or favorable evidence; but *something* seems to have increased in either case—we have a more substantial basis on which to rest our conclusion. . . New evidence will sometimes decrease the probability of [the hypothesis] but will always increase its "weight." (Keynes 1921, 77, qtd. in Joyce 2005, 159)

The intuitive idea, then, is that the weight of the evidence is a function of how much total evidence informs one's credence in P at t, irrespective of the value of $cr_t(P)$. It is tricky to apply this idea at an intuitive level to COIN CASE and CARD CASE. One is tempted to say that reason belief in H is justified after a reasonably high number of consecutive heads trials in COIN CASE is that one has *weighty* evidence for H. And, indeed, it is clear that in COIN CASE, each consecutive heads trial adds to the weight of the evidence for H (and for B). But there is also a limit to how weighty the evidence can be; you cannot get weightier evidence if you cannot get any more relevant evidence. So, in COIN CASE, a single tails trial after any

[18]I am ignoring the third property of evidence that Joyce (2005) appeals to, namely, specificity. The reason for this is that specificity can only be measured in a generalization of the normal Bayesian framework that treats credences as possibly imprecise and models credal states with sets of probability functions (see § 2.3.4). I do think such a generalization is a good idea in general, but one of the virtues of COIN CASE and CARD CASE is that we can work with the simpler Bayesian framework because the evidence is maximally specific—a perfectly precise credence is mandated in each case. So a difference of specificity is inadequate to explain why we should believe H in COIN CASE but not $\neg A$ in card.

number of heads trials makes one's credence in H $\frac{1}{2}$ forevermore, regardless of the results of future trials—one's evidence cannot get weightier, at that point. So the explanation of why belief in H is justified in COIN CASE but belief in $\neg A$ is not justified in CARD CASE will not be that the evidence is weightier in COIN CASE. For CARD CASE *starts* with maximally weighty evidence—we could draw cards from the deck and replace and reshuffle, and the results of these trials would have no effect on our credence in A. Given the setup of the case, we are already certain of a chance hypothesis that screens off the relevance of such evidence. So in fact, we might want to say that the evidence in CARD CASE for $\neg A$ is actually weightier than the evidence for H in COIN CASE, no matter how many consecutive heads trials we get. And on the proposed measure of weight in Joyce (2005), this is indeed correct.

Joyce (2005, 165) admits that "no satisfactory measure of the weight of evidence has yet been devised." Nevertheless, he attempts to provide a partial measure of weight for cases in which one's credence in P is mediated by one's credence in hypotheses about the chance of P. The motivation for this measure is the idea that we can judge the weight of a body of evidence in terms of the resilience of credences based on that evidence under various assumptions of possible additional evidence.

> [T]he point, made persuasively by Skyrms (1980), [is] that the weight of the evidence for a proposition X often manifests itself not in X's unconditional credence, but in the *resilience* of this credence conditional on various potential data sequences. A person's credence for X is resilient with respect to datum E to the extent that her credence for X given E remains close to her unconditional credence for X. Note that resilience is defined relative to a specific item of data: a person'a belief about X may be resilient relative to one kind of data, but unstable with respect to another. That said, it is usually the case that the greater volume of data a person has for a hypothesis the more resilient her credence tends to be across a wide range of additional data. (Joyce, 2005, 161)

Let us consider a simple CARD CASE in which we have a deck of 66 $(2 + 2^{5+1})$ cards. The chance of drawing an Ace (A) is $\frac{1}{66}$. Suppose we draw a card, and replace it and reshuffle. Do this as many times as you like. Let P be the evidence you have just from the set-up of CARD CASE, and let Q be the evidence from these trials. Then you should have $cr_P(A) = cr_P(A|Q) = \frac{1}{66}$, i.e., your credence in A should have changed not at all. Without going into the technical details, Joyce's measure of weight reflects this; the quantity $w(A, P)$, which measures the weight of evidence P for A (the lower the weightier), will be zero, and so will $w(A, PQ)$.[19] So if smaller $w(X, E)$ is supposed to mean that E is weightier evidence for X, P and PQ are both maximally weighty evidence for A. In fact, just about any evidence for A[20] that includes P will be maximally weighty, precisely because just about any evidence is irrelevant given our certainty about the chances.

This is a slightly counterintuive artifact of the measurement, but an acceptable one. It is counterintuitive because one might expect maximally weighty evidence to be *decisive*, defined as follows.

> Decisive Evidence. Evidence E is decisive evidence for P, for a subject S at time t, if and only if, given that S is certain of E, it is epistemically impermissible for S to suspend judgment about P at t.

Clearly it is possible to have maximally weighty evidence that supports a credence of $\frac{1}{2}$ in a proposition, so decisive evidence is false. This is nevertheless an acceptable (or even desirable) consequence for a measure of the weight of the evidence for a proposition. Maximally weighty evidence is the best we can expect to do. But there may well be no reason to expect that we will be able to decide every question that we might like to. While general skepticism is problematic, the possibility that certain propositions might be unknowable is unobjectionable.

We can apply this measure to coin case and card case in order to explain why belief in H is justified in coin case, but belief in $\neg A$ is not justified in card case. The first step is to consider the weight of the evidence not for the proposition that the next flip will be heads or that an ace will *not* be drawn, but for the corresponding "skeptical hypotheses" that the next flip will be tails or that an ace *will* be drawn (T and A). These are the salient mistakes that might undermine our justification to believe. Because we start with maximally weighty evidence for A, we cannot reasonably expect to lower our credence in A. It will remain a salient mistake no matter what evidence we get. But although after n consecutive heads flips our credence in T will be the same as our credence in A for a deck of

[19]The measure is defined as follows: for all x such that one gives some credence to the proposition that the chance of X is x, $w(X, E) = \sum_x |cr(\langle ch(X) = x\rangle|E) \cdot (x - cr(X|E))^2 - cr(\langle ch(X) = x\rangle) \cdot (x - cr(X))^2|$. In the case at hand, there is only one relevant chance hypothesis, and we are certain of it: $ch(A|P) = \frac{1}{66}$. So since we will have $cr_P(A) = cr_P(A|Q) = \frac{1}{66}$, $w(A, P) = w(A, PQ) = 0$.

[20]Or $\neg A$—weight is supposed to be independent of balance, and indeed this measure of weight has the property that $w(X, E) = w(\neg X, E)\ \forall X, E$.

$2 + 2^{n+1}$ cards, the evidence against T is not maximally weighty. Moreover, we will expect to get *more* evidence against T if we keep flipping the coin, since $cr_n(H) > cr_n(T)$. Finally, we do not expect to get maximally weighty evidence against T—we know, given the setup, that this is impossible.

These differences in the evidence against A and the evidence against T are relevant to the question whether we should believe H and $\neg A$, and can explain the intuitive difference in our judgments about COIN CASE and CARD CASE. My claim is that the difference between these cases is that $\neg A$ is unknowable in CARD CASE, because given the setup, we know that we are not safe from A, but H is possibly knowable, since for all we know, the coin is biased and we are safe from $\neg HF$. Moreover, after several consecutive heads flips, we have relatively weighty evidence that we are safe. Even if I have gotten the explanation somewhat wrong (perhaps safety is not a necessary condition on knowledge), I have still identified a difference between the cases that a Bayesian epistemologist can appeal to in explaining the differece, where no such difference is available if we focus solely on the balance of the evidence.

The diagnosis of why THRESHOLD and preference invariance accounts of the relationship between credence and belief fail for these cases, then, is that they make belief entirely a matter of the balance of the evidence (and the subject's conditional preferences) and are incapable of taking weight into account. JCB is false, but the modified JCB* may not be. So although COIN CASE and CARD CASE taken together can constitute a counterexample to many reductive accounts of the relationship between credence and belief, they do not undermine the probability first approach to epistemology generally.

In the following chapters, therefore, I will adopt a probability first approach. Given that I do not have a decisive argument that a probability first approach is correct, those readers who do not share my conviction that it is likely to be correct should regard my arguments as showing what those probability first epistemologists should think if they wish to remain consistent with their probablistic commitments, rather than what just anyone should think. I regret that I cannot do more than that here, but my hope is that the conditional conclusions are sufficiently interesting in any case, given the prevalence of probabilistic arguments in the current epistemological literature.

2.0 SKEPTICAL PROBLEMS FOR INFERENTIAL KNOWLEDGE

2.1 INTRODUCTION

Views about *how* we are justified in believing the conclusions of our inferences have consequences for *when* we can or must be justified in various other beliefs. This chapter explores the interaction of views in traditional epistemology about how our inferential beliefs are justified with views in Bayesian epistemology about how our credences should change over time.

§ 2.2 presents a schematic argument that a skeptic might exploit about *when* certain beliefs may be justified in order to motivate a division of logical space into different anti-skeptical views that deny the premises of this argument. After defending the second premise of this argument in §2.2.1, I discuss several different accounts of what is wrong with the first premise: the inferential internalism of Fumerton (1995), the dogmatism of Pryor (2000), and the rationalist liberalism of Silins (2007).

After canvassing some views about *how* beliefs are justified, in order to sort them according to what they say about *when*, I attempt to bring the probabilistic tools of Bayesian epistemology to bear on the question. Orthodox Bayesians have a story about how credences ought to evolve over time. Given some highly plausible assumptions about the relation between belief and credence, the orthodox story entails that the dogmatist's views about how anti-skeptical beliefs are justified entail impossible consequences about when they are. I present this argument, which is originally due to White (2006), in § 2.3. In § 2.3.1, I argue that the assumption that credences should change by conditionalization, in line with the Bayesian orthodoxy, does undermine dogmatism.

However, I think there are good independent reasons to reject the orthodox Bayesian line about the evolution of credences. §2.3.3 and §2.3.4 explore two different versions of reformed Bayesianism that might be of use to the dogmatist. I conclude that the Bayesian argument against dogmatism is not decisive. However, all the formal views compatible with dogmatism (and some that are not) require us to introduce some source of constraint on the evolution of credences other than propositions made certain by experience and conditionalization on those propositions, and this makes dogmatism unnecessary, since this additional constraint can suffice to independently justify $E \supset H$ when we are in a position to gain justification for H by inference from E. Moreover, the one formal view that is compatible with dogmatism is poorly motivated independently of a desire for a model compatible with dogmatism; it is main alleged virtue is not unique to it. I conclude that rationalist liberalism is better supported by probabilistic considerations.

2.2 A SCHEMATIC ARGUMENT FOR SKEPTICISM

There are many ways of arguing for skepticism about some belief or beliefs; generally speaking, a skeptical argument that one is not justified in believing a hypothesis H proceeds by showing that one is not justified in believing some other proposition Q on which the justification for H allegedly depends. More general skeptical arguments find a single proposition on which the justification of beliefs of a certain type allegedly depend, and argue that we lack justification for it. For example, one might argue (as Descartes may have) that my perceptual and even mathematical beliefs are justified only if I have a justified belief in the proposition that a non-deceiving god exists, or (as Hume may have argued) that my beliefs about causal relationships are justified only if I have a justified belief that the future will resemble the past. In both cases, the skeptic will go on to argue that the relevant proposition is not or cannot be justified.

The specific domain of beliefs that I am primarily concerned with are beliefs supported by non-demonstrative inference from empirical evidence; roughly the class of beliefs that Hume was concerned with (inferential beliefs, for short).

As before, we can distinguish, between a proposition H being justified for a subject S

and S's having a well-founded belief in H. The basic idea is articulated by Feldman and Conee (2004, p. 83):

> Doxastic attitude d toward proposition H is epistemically justified for S at t if and only if having d toward H fits the evidence S has at t.

Well-foundedness differs from this sort of propositional justification because S may not in fact believe what her evidence supports, or may believe H on the basis of evidence that does not support it, or may believe H on the basis of evidence that *does* support it but which is defeated by other misleading evidence she is improperly ignoring. A well-founded belief satisfies these additional criteria. It is the notion of propositional justification that I will be primarily concerned with, not the property of well-foundedness.

One need not find a single proposition on which every non-demonstrative inference depends in order to argue for a general skeptical position with respect to this sort of justification. It is enough if, whenever H is inferred by S from some propositional evidence E that does not entail H, there is some proposition Q which is such that H is not justified for S and, if Q is unjustified for S, the inference does not confer justification. Say that some evidence E *justifies* a hypothesis H for S at t just in case S has evidence E, S's total evidence permits an attitude of belief in H, and S's total evidence minus E would not permit an attitude of belief in H. Here is a form of schematic skeptical argument for any proposition H and evidence E from which it is inferred.

1. $E \supset H$ is not justified for S at t.
2. If $E \supset H$ is not justified for S at t, then E does not justify H for S at t.
3. S is not justified in believing H at t on the basis of an inference from E.

I intend the "$\supset$" here to be read as the material conditional. Why choose this particular material conditional as the relevant proposition for the skeptic to target? $E \supset H$ has the interesting property of being entailed by H and entailing H jointly with the evidence E. Assuming a plausible principle of epistemic closure, this means that justification for H on the basis of knowledge of some evidence E must come along with a justification to believe $E \supset H$, and similarly if one knew that $E \supset H$, then justification to believe E would come along with a justification to believe H. It is important to note that this skeptical argument makes no assumptions about whether one must be justified in believing $E \supset H$ *prior* to the time at which she is justified in believing H on the basis of E. The skeptic advancing this argument claims that justification to believe $E \supset H$ is a necessary condition of having inferential justification to believe H on the basis of E; she need not claim that it is a necessary pre-condition. For all the argument says, it would be enough if H justified $E \supset H$ for S, rather than the other way around.

While the schematic skeptical argument above does not have any real plausibility in the abstract, this feature will allow us to isolate a few different ways of being an anti-skeptic. In particular, the anti-skeptic may deny the first premise by claiming that $E \supset H$ is justified for S at t, whether independently of and antecedently to S's acquiring justification for H on the basis of E (so that, perhaps, $E \supset H$ ends up being part of the inferential basis for H) or not. Alternatively, it may be possible to resist the skeptical argument by denying the second premise; i.e., by supposing that despite these entailment relations, it is just not necessary for $E \supset H$ to be justified for S when S acquires justification to believe H on the basis of E. There are many ways this might go, and has gone in the literature. Some prominent views that offer various explanations of what is wrong with this schematic argument are canvassed below.

2.2.1 EPISTEMIC CLOSURE PRINCIPLES

Probably the least popular view of what is wrong with skeptical arguments like the one above is that premise (2) is false. Since, if S *is* justified in believing H at t because S infers H from E but (1) is true, so that S is *not* justified in believing in believing $E \supset H$ at t, then the following principle of closure of justification must not be universally true.

> Justification Closure. If, at the same time, S is justified in believing P and S knows that P entails Q, S has a justification for believing Q at that time.

For we can assume knowledge that H entails $E \supset H$. While perhaps not everyone knows this, if the skeptical argument is to be generally applicable it must apply to those of us who do. So (2) and $\neg$(3) jointly entail $\neg$(1), if we assume justification closure.

The related principle of epistemic closure for knowledge is similar, and has more often come under attack.

Knowledge Closure. If, at the same time, S knows that P and S knows that P entails Q, S is in a position to know Q at that time.

S may not actually know that Q, for S may not make the relevant inference, or may disbelieve Q.

Dretske (1970, 2005) and Nozick (1981) deny knowledge closure. Both explicitly frame skeptical arguments as relying on closure. Here is Nozick:

> In taking the "short step," the skeptic assumes that if S knows that p and he knows that p entails q, then he also knows that q. In the terminology of the logicians, the skeptic assumes that knowledge is closed under known logical implication; that the operation of moving from something known to something else known to be entailed by it does not take outside of the (closed) area of knowledge. He intends, of course, to work things backwards, arguing that since the person does not know q, assuming (at least for the purposes of argument) that he does know that p entails q, it follows that he does not know that p. For if he did know that p, he would also know that q, which he does not.

Nozick and Dretske both think that the first premise of the skeptic's argument, which has it that one does not know some proposition entailed by what one thinks one does know, is often true, so that denying closure is our only hope for evading skepticism about all kinds of everyday knowledge.

> Granted, the [skeptical] hypothesis (if we may call it that) is not very plausible.... But the question here is not whether this alternative is plausible, but whether *you know* that this alternative hypothesis is false. I do not think you do. (Dretske, 1970)

Dretske and Nozick support their claims that we do not know that skeptical hypotheses are false with appeals to intuition, rather than arguments.

> The skeptic asserts that we do not know his possibilities do not obtain, and he is right. Attempts to avoid skepticism by claiming we do know these things are bound to fail. The skeptic's possibilities make us uneasy because, as we deeply realize, we do not know they do not obtain; it is not surprising that attempts to show we do know these things leave us suspicious, strike us even as bad faith. (Nozick, 1981)

Dretske (2005) gives a name to this class of propositions that we allegedly "deeply realize" we do not know—he calls them "metaphysically heavyweight." Beyond the clear intention that typical skeptical hypotheses and the few examples Dretske discusses are metaphysically heavyweight, this class of propositions is not clearly delineated.[1] I may say that I do not

[1]The nearest thing to a definition of "metaphysically heavyweight" in Dretske (2005) is the claim that a proposition Q that is entailed by a known proposition P is metaphysically heavyweight if Q is not known

share the intuition that attempts to show that we know that skeptical hypotheses can be known to be false are in bad faith. If such persistence in the face of a deep mystery about how to proceed is bad faith, then I do not see how the attempt to preserve our everyday, common-sense knowledge in the face of skepticism is any better. In fact, I find the arguments of Hawthorne (2005) persuasive in showing that there is no deep distinction between "metaphysically heavyweight" and "everyday" propositions, so that this intuitive judgment does not stand up to scrutiny. There is nothing metaphysical and mysterious about the claim, for example, that a red-seeming table is in fact white but illuminated by red lights. But we fail to know that this "skeptical hypothesis" is false by both Dretske's and Nozick's lights.

To the extent that I can work myself into sharing the intuitions about skepticism that Dretske and Nozick express (I have, certainly, found the the difficulty of making any progress towards defending the claim that I know that I am not a brain in a vat extraordinarily daunting), it is also worth noting that these intuitive judgments are most gripping in skeptical arguments that target knowledge, rather than justified belief. I may not *know* that I am not being deceived by an evil demon, since I can conceive of the possibility that I am wrong in that belief. But I certainly feel that I may believe, with some justification, that I am not being deceived by an evil demon. As Dretske grants, skeptical hypotheses are not typically particularly plausible. But if it is possible for me to have justification to believe that skeptical hypotheses are false, and if they *are*, in fact, false, then I have a hard time seeing what, in principle, the problem with knowing that they are false is. It is surely true that we cannot be absolutely certain that skeptical hypotheses do not obtain, but this just as naturally regarded as a reason to be *fallibilists* about knowledge rather than a reason to think that we cannot know. There are problems with interpreting fallibilism in such a way that it is compatible with the truism that knowing that P is incompatible with it is being epistemically possible that $\neg P$; as Lewis (1996) argues, the following sentences are infelicitous:

I know that P, but there is a chance that not-P.
I know that P, but it is possible that not-P.

on the basis of the same method as P; paradigmatically, consequences of perceptually-based beliefs that are not themselves known by perception are metaphysically heavyweight. There is clearly no reason to think that we do not know propositions which are metaphysically heavyweight only in this sense, unless perhaps we already accept that knowledge is not deductively closed.

This may seem to conflict with a definition of fallibilism such as the one provided by Fantl and McGrath (2009).

> FALLIBILISM. Possibly there is a non-zero epistemic chance for S of $\neg P$ but S knows that P.

But Hawthorne and Lasonen-Aarnio (2009) and Williamson (2009) have argued that non-zero objective chance that P, and even high non-zero objective chance, does not entail that there is a close risk of P, in the sense of "close" relevant to knowledge. Either the same is true of epistemic chance, or whenever there is no close risk of $\neg P$, then there is no epistemic chance that $\neg P$. Either way, we might try to explain the infelicity of the combining knowledge ascriptions with statements of the possibility/chance of error in Gricean terms; perhaps "there is a chance that not-P" typically communicates that the risk of not-P is close in the relevant sense, even if there is no entailment from the fact that there is an epistemic chance that not-P to the fact that one does not know that P.[2]

Moreover, it is clear that an attitude of absolute certainty in what one knows is not an epistemic virtue even in an ideally rational human being.[3] And there is an unobjectionable way of describing a proper attitude towards our own beliefs that a requirement of absolute certainty for knowledge appears to conflict with, at least at first glance.

> For our purposes let us understand fallibilism as the view that we should hold every belief, no matter how strongly it is supported, in an open minded spirit which acknowledges the possibility that future evidence may require us to abandon it. (Casullo, 1988)

Even if we only accept this truism, it removes much of the intuitive force of the skeptic's premise (1). For this attitude requires us to acknowledge that there are possible reasons to doubt that do not undermine justification—not all doubts are reasonable. I do not claim to be able to establish fallibilism about knowledge or justification here, or even to know how to precisely formulate an adequate statement of fallibilism. Nevertheless, fallibilism is intuitively at least as plausible as skepticism.[4] Since the claim that the skeptic's premise (1) is intuitively obvious is the main support for denying KNOWLEDGE CLOSURE, and both

[2] Cf. Dougherty and Rysiew (2009); Fantl and McGrath (2009).

[3] Cf. Christensen (2007a).

[4] Despite arguing that fallibilism seems mad, Lewis goes on to defend fallibilism on the grounds that it is the "less intrusive madness" (Lewis, 1996, 550).

KNOWLEDGE CLOSURE and JUSTIFICATION CLOSURE are themselves highly intuitive principles, I will assume in what follows that knowledge and justification are closed under known entailment. Only if it can be shown that there is really no hope of a straightforward response to the skeptic will it be appealing to deny closure.

2.2.2 INFERENTIAL INTERNALISM

My skeptical argument above is quite similar to the variety of skeptical argument that Fumerton (1995) is primarily concerned with. Fumerton claims that the skeptic implicitly assumes the following principle.

> PRINCIPLE OF INFERENTIAL JUSTIFICATION. To be justified in believing one proposition P on the basis of another proposition E, one must be (1) justified in believing E and (2) justified in believing that E makes probable P. (Fumerton, 1995, 36)

Fumerton prefers the proposition that E makes probable P as the target of skeptical attack, rather than the associated material conditional, and offers the following explanation for this preference.

> [O]ne might claim that to be justified in believing P on the basis of E, one must be (1) justified in believing E and (2) justified in believing that if E then P. But one must immediately inquire as to the interpretation of the conditional, if E then P. One thing seems obvious. If one is trying to understand how one could be justified in believing P *through* or *by* justifiably believing the conjunction of E and if E then P, one should not interpret the conditional as a contingent, truth-functionally complex conditional of material implication that is not made true by a *connection* between E and P. . . .
> The only conditional that one could plausibly appeal to in a principle of inferential justification is one that asserts a *connection* between E and P. But what kind of connection should we view the conditional as asserting? Entailment is one candidate, but we can surely justifiably believe P on the basis of E even when E does not entail the truth of P. What we seem to need is the kind of connection that holds between the premises and conclusion of a good argument. When a deductively invalid argument is good, the premises must make probable the conclusion. To be justified in believing P on the basis of E, one must be justified in believing that the inference from E to P is legitimate. To justifiably believe that an inference from E to P is legitimate just *is* to justifiably believe that E makes probable P (where E's entailing P can be thought of as the upper limit of making probable). (Fumerton, 1995, 88)

I present Fumerton's inferential internalism here because it is a good example of the sort of view of inferential justification that Silins (2007) and Pryor (2004) would call "conservative."

Silins' statement of conservatism is defined for justification on the basis of visual experience only, but it is not hard to state an analogous version that applies to inferential justification.

> Conservatism. Whenever S infers some hypothesis H from evidence E against background knowledge K, where EK does not logically entail H,
> (i) evidence E does not give S immediate justification to believe that H, i.e., what justifies S in believing H includes her having some independent justification to believe other propositions,
> and in particular, (ii) what makes E justify S in believing that H includes her having independent justification to believe $\neg(E\neg H)$, i.e., $E \supset H$.

Fumerton's principle of inferential justification is conservative because having justification to believe that E makes probable H provides one with independent justification to believe $E \supset H$.

Fumerton is proposing an account of *how* one might be justified in believing H on the basis of an inference from E. He claims that if one is not justified in believing that E makes probable H because one "sees" the connection between E and H, then one cannot be justified in believing H on the basis of an inference from E. Fumerton spells out this metaphor of seeing in terms of direct a priori insight into epistemic relationships between propositions. These relationships are meant to be more robust than the relationship that is described by the material conditional $E \supset H$, which does not support any counterfactual judgments. H may very well be false at *every* close world in which E is true, despite the actual truth of the material conditional. Fumerton clearly wants a relationship that one can have justification to believe obtains independently of and antecedently to acquiring justification for H on the basis of E, and which *explains* why E justifies belief in H. The material conditional cannot do this explanatory work because it may very well be true even though there is in fact no connection between E and H, or worse, it may even be true despite the fact that E would justify belief in $\neg H$ in the absence of any other reason to believe H.

However, as the passage above makes clear, "inferential externalist" views (which deny the need for justification to believe that E makes probable H) may also be conservative. Assuming closure, one necessarily has justification for $E \supset H$ whenever one has justification for H, whatever the source of that justification (and even if E is logically and probabilistically independent of H). So one might have justification for believing $E \supset H$ because one has justification for believing H, even though one's priors are such that $p(H|E) = p(H)$ (i.e., one previously took E and H to be probabilistically independent); the probability will be high after learning H only because the probability of H has increased. It would be bizarre, under such circumstances, to claim that E makes probable H in Fumerton's sense.

However, while this is possible, having an independent justification to believe that $E \supset H$ when one has credences such that $cr(E) < 1$ entails that one's credences treat H and E as probabilistically *inter*dependent. The special case in which $cr(E) = 1$ is just the standard Bayesian problem of old evidence; anyone who is already certain that E will have credences that treat E as probabilistically independent of every other logically consistent proposition. And if $cr(E) < 1$, then suppose that one has credences that treat H and E as probabilistically independent. It is easy to see that $cr(E \supset H) = cr(\neg E) + cr(H|E)$, since $cr(E \supset H) = 1 - cr(E\neg H)$. But if H and E are independent, then $cr(H|E) = cr(H)$. So then one's credence in $E \supset H$ is precisely the sum of one's credence in $\neg E$ and one's credence in H. If one thereafter gains independent evidence that confirms $E \supset H$ but *not* by way of confirming either $\neg E$ or H, necessarily one's credence distribution should change so that $cr(EH)$ is confirmed. But if $cr(E) < 1$, then since $cr(H|E) = \frac{cr(HE)}{cr(E)}$ and $cr(H) = cr(EH) + cr(\neg EH)$, $cr(H|E)$ will be raised to a degree greater than $cr(H)$. So it will no longer be true that $cr(H|E) = cr(H)$, i.e., one's new credence distribution will no longer treat H and E as independent. Or, to put this result more succinctly, for all probability functions p, $p(H|E \mathbin{\&} E \supset H) \neq p(H|E \supset H)$ unless $p(E) = 1$.

So, although one might learn $E \supset H$ by learning H, rather than the other way around, if one *does* have independent justification for $E \supset H$, i.e., if one has reason to give more credence to $E \supset H$ than one has just based on one's credence in E and H, then one is also justified in having a credence distribution that treats E as evidence that confirms H. For this reason, I think the debate over CONSERVATISM justifiably focuses on the material conditional. While justified belief in $E \supset H$ may not support counterfactual claims about the relationship between E and H, it does have consequences for what one's credence distribution must look like—consequences that will entail, together with PROBABILISM and CONDITIONALIZATION, further consequences about how learning E should affect one's credence in H. And the converse holds as well; claims about how learning E should affect one's credence in H will have consequences for what one's *prior* credence in $E \supset H$ may be.

As pointed out above, denying (1) of my schematic skeptical argument need not commit one to CONSERVATISM, and recently "liberal" views of perceptual justification have gained currency—particularly the "dogmatism" of Pryor (2000) and the "rationalist liberalism" of Silins (2007). Analogous views for inferential justification would have in common the following thesis.

> LIBERALISM. It is not the case that: whenever S gains justification for some hypothesis H by inferring it from evidence E against background knowledge K, where E does not entail H, what justifies S in believing H includes her having independent justification to believe some other proposition.

We can say that an attitude of belief in $E \supset H$ is justified for S *independently* of some evidence E if S would be justified in believing that $E \supset H$ given only her background knowledge K.

However, given JUSTIFICATION CLOSURE, it is still necessary that S have justification to believe $E \supset H$ whenever S has justification to believe H. The dispute between rationalist liberals and dogmatists is over whether that justification is independent of her justification for H or not. Given that LIBERALISM says that the justification for $E \supset H$ when one infers H from E need not be independent, the view that it is not independent in some cases may seem like the more natural form of liberalism. This is DOGMATISM, which will be developed in the next section, before we turn to RATIONALISM. Then, in §2.3, I will present an argument (originally due to White (2006)) that attempts to exploit the probabilistic relations discussed above to show that DOGMATISM is false.

2.2.3 DOGMATISM

I have been talking about inferential justification, but Pryor (2000) defends dogmatism about perceptual justification. I am not sure that dogmatism about perceptual justification is wrong, and will be arguing against what I shall call "inferential dogmatism." But first, here is how Pryor describes his perceptual dogmatism.

> The dogmatist about perceptual justification says that when it perceptually seems to you as if p is the case, you have a kind of justification for believing p that does not presuppose or rest on your justification for anything else, which could be cited in an argument (even an ampliative argument) for p. To have this justification for believing p, you need only

> have an experience that represents p as being the case. No further awareness or reflection or background beliefs are required. Of course, other beliefs you have might defeat or undermine this justification. But no other beliefs are required for it to be in place. (Pryor, 2000, 519)

In particular, the perceptual dogmatist supposes that in order to acquire perceptual justification for believing P, one need not *first* or *independently* have justification for believing that one's experience is veridical, that skeptical hypotheses do not obtain, that one's experience makes probable P, or anything of the sort that the skeptic will attempt to show one does not have justification for. As noted above, this has the somewhat counterintuitive consequence, given closure, of endorsing Moore's famous proof of the existence of an external world. For if I can gain perceptual justification for believing that I have hands just by having it perceptually seem to me as if I have hands, and I do not *already* have justification for believing that there is an external world, but I *do* know that if there are hands, then there is an external world, I can infer and thereby gain justification for believing that there is an external world. The skeptic will, of course, complain that this begs the question, but the dogmatist does not mind; she seeks only a reflective reason for thinking that premise (1) of the skeptic's argument is false that will satisfy us non-skeptics, not one that will be dialectically effective against the skeptic.[5]

Inferential dogmatism, then, should similarly be a combination of LIBERALISM about inferential justification and a Moorean thesis that both accepts that one has justification for believing $E \supset H$ whenever one has justification for believing H and denies that this justification need be independent of one's justification to believe H. The Moorean thesis will be something like:

> DOGMATISM. When S gains justification for some hypothesis H by inferring it from evidence E against background knowledge K, S is justified in believing $E \supset H$, but possibly S was not justified in believing $E \supset H$ prior to learning E (on the basis of K alone).

Inferential dogmatism is potentially appealing as a response to skepticism because of the disquiet that moved the closure-deniers. If we worry that Humean skeptical arguments undermine whatever reasons we might have to believe that there are real epistemic connections

[5] This is not to say that there are not serious worries about whether Moore's argument is question-begging in a deeper sense; see Pryor (2004) for a discussion of precisely what the dogmatist will and will not admit is wrong with Moore's argument from a dialectical standpoint.

that we can "see" in something like Fumerton's sense between the evidence we get from experience and the hypotheses we use to explain that evidence in terms of unseen causes or other entities of which we have no direct experience, then we might be strongly motivated to want an epistemology of inference that allows us to justify our beliefs in the unobservable without having to defend on any independent grounds the idea that there is some epistemic connection between the evidence and our hypotheses. How do we know that H? Well, we have some evidence—E. How do we know that E really supports H? Perhaps we do not, but if we know H, we know that $E \supset H$, even if we cannot justify this conditional by appealing to any uncontroversial way in which E provides evidential support for H. That can come later, and be a task for epistemologists, if it is possible at all.

Dogmatism (both varieties) may also be motivated by the thought that it better represents our pre-theoretical intuitions about what justification requires.

> For a large class of propositions, like the proposition that there are hands, it is intuitively very natural to think that having an experience as of that proposition justifies one in believing that proposition to be true. What is more, one's justification here does not seem to depend on any complicated justifying argument. An experience as of there being hands seems to justify one in believing that there are hands in a perfectly straightforward and immediate way. When asked, "What justifies you in believing there are hands?" one is likely to respond, "I can *simply see* that there are hands." One might be wrong: one might not really be seeing a hand. But it seems like *the mere fact* that one has a visual experience of that phenomenological sort is enough to make it reasonable for one to believe that there are hands. No *premises* about the character of one's experience—or any other sophisticated assumptions—seem to be needed. (Pryor, 2000, 536)

Similarly, though of course any inferentially-supported belief *will* involve a potential justifying argument, that argument *will not* typically include as a premise the claim that the premises of the inference support the conclusion, or we will end up in Achilles and the Tortoise territory. Of course, just as Achilles knows that if P, then Q and *can* assert this conditional, though he need not in justifying his belief that Q, it may be that a sufficiently reflective or sophisticated reasoner will, generally, be able to add to the evidence that supports her conclusion the claim that the evidence supports the conclusion. The inferential dogmatist merely denies that she *must* be able to do so *before* we will grant that her evidence really does support her conclusion.

Dogmatism has its attractions, but it also has its detractors. In §2.3, I will present what I take to be the biggest problem for dogmatism. First, though, I want to visit one last view that may be thought of as occupying a sort of middle ground between inferential internalism and dogmatism.

2.2.4 RATIONALIST LIBERALISM

As Silins (2007) has taken pains to show, LIBERALISM does not entail DOGMATISM. It is possible to hold on to LIBERALISM without endorsing Moorean reasoning as a way of gaining justification.

> The central Liberal claim is that some of our perceptual beliefs are non-inferentially justified, that is, justified in a way that does not involve our having independent justification against this or that skeptical hypothesis. The Liberal does not yet say that any of our perceptual beliefs are justified despite our lack of independent justification against skeptical hypotheses. (Silins, 2007, 130)

The statment of LIBERALISM above includes the somewhat vague notion of what is "included" in whatever justifies an inferentially-justified belief. In motivating DOGMATISM, as noted above, it is common to appeal to the fact that in defending an inferentially-justified belief one typically does not cite anything like a belief that inductive inference is reliable, or even that this evidence evidentially supports that hypothesis, or other such "higher-level" epistemic beliefs—one simply appeals to the evidence. Even if we take this to heart and suggest that what is included in whatever justifies belief in H is just whatever one would appeal to in a complete, satisfying explanation of why one believes H, it does not follow that there is not much more that one has propositional justification to believe when one has such an inferentially-justified belief. It may just be that this extra stuff is not "included in" what justifies H. That this can happen is easily demonstrated with some examples in which it is either intuitively clear that the extra stuff plays no role in supporting H, or at least the claim that it does not play a supporting role has been attested to in the philosophical literature.

> First, consider the contingent proposition that I exist. Presumably, whenever I have perceptual justification to believe that I have hands, I have a cogito-style independent justification to believe that I exist. But we should not infer from this that I am not immediately justified

> in believing that I have hands. Second, consider testimony and contingent propositions I do not always have justification to believe. It may well be that, whenever a source's testimony justifies me in believing that p, my experience plays some role in giving me justification to believe that the source testified that p. We cannot yet infer that my testimonial justification is to be explained in terms of my experiential justification (Burge, 1993). Third, consider introspection and contingent propositions I do not always have justification to believe. It might be that, whenever my experience justifies me in believing that p, something other than experience (inner sense, perhaps) justifies me in believing that I have the experience. We cannot yet infer from this that my experience does not immediately justify me in my belief. In sum, it seems perfectly possible for me to gain immediate justification to believe P from source S even if, whenever I gain immediate justification to believe P from S, I have independent reason to believe Q.

Similarly, it is open to a liberal about inferential justification to deny DOGMATISM and endorse the following principle.

> RATIONALISM. Whenever S gains justification for some hypothesis H by inferring it from evidence E against background knowledge K, S is justified in believing $E \supset H$ independently of E.

It can be very hard to see how RATIONALISM and LIBERALISM are compatible, at least on a probability first story like the one I want to tell. If I *must* always be justified in believing $E \supset H$ when I infer H from E, and if what justifies me in believing *any* proposition is just having a reasonable credence distribution (and perhaps preferences and other non-epistemic features) that entails that I believe it, then surely the same credence distribution will justify both $E \supset H$ and H, given E. I suspect a defender of rationalist LIBERALISM will say that surely it is not the case that *every* proposition I am justified in believing is "included in" my justification for believing H when I infer it from E. For example, it is no part of my justification for believing that all emeralds are green that the square root of two is irrational, and I am justified in believing the latter proposition on the basis of the permissibility of the same credence distribution that makes it rational for me to believe that all emeralds are green on the basis of evidence about the colors of observed emeralds. Nor can we simply say that I might be justified in making this inference even if I were not justified in believing that the square root of two is irrational, for on a typical probabilistic story, I am *always* justified in believing such mathematical claims. Insofar as the intuitive notion of one's justification for a proposition being "included in" what justifies belief in another can be made more precise, it is not obvious how to do so using facts about one's credence distribution.

I suspect we could go some way to specifying what we mean here by appeal to *non*-ideally rational agents, or ideally rational agents in compromised epistemic positions. Given that I am not ideally rational, I nevertheless approximate rationality to some degree. I might (approximately) rationally believe H on the basis of E while failing to believe some complex logical theorem that I rationally ought to give credence 1, though I couldn't even approximately rationally believe H on the basis of E while failing to believe E. We might then say that justification to believe Q is "included in" the justification for a belief that P if one could not even approximately rationally believe that P without believing that Q. I do not have strong intuitions about how to manage this sort of "approximate rationality" talk; if something like this is right, then a defender of rationalist LIBERALISM will have to say that I might fail to believe $E \supset H$, despite having justification for it, while nevertheless approximately rationally believing H on the basis of E. Perhaps the rationalist liberal will reject this attempt to help. But some distinction needs to be made here if LIBERALISM is not to collapse into DOGMATISM. While I do not know how to draw the line around what is "included in" one's justification to believe H, I am open to the idea that there is a good way of spelling out this distinction, so that LIBERALISM need not be incompatible with RATIONALISM.[6]

In what follows, I will argue that the Bayesian argument against DOGMATISM is successful if we adopt fairly orthodox Bayesian assumptions, but there is no argument from the same assumptions against a combination of RATIONALISM and LIBERALISM, on the assumption that such a combination makes sense. However, the Bayesian argument is not decisive, since there is good independent reason to reject some of the orthodox Bayesian assumptions on which it depends (even if we want to endorse the broadly-defined program of "Bayesian"

[6]It may be that this distinction is also needed to explain cases like those that that Willenken (2011) uses to attack what he calls "robust conservatism." In his cases, a subject S has what Christensen (2010) calls "higher-order evidence" that her ability to judge whether $E \supset H$ is compromised, even as she reasonably infers H from E (as her ability to perform this inference is not compromised). For inquirers in such a compromised position, it may well be impossible to live up to our epistemic ideals. One might say that what this case shows is that justification to believe $E \supset H$ is not "included in" what justifies H, though ordinarily (except in this sort of weird case) one has independent justification to believe $E \supset H$ whenever one reasonably infers H from E. One may even yet have propositional justification to believe $E \supset H$ in such situations; what cases like this suggest is that it the maxim: "believe what one has propositional justification to believe" may not always be good advice, and may conflict with other epistemic ideals (see also the problem of logical omniscience).

epistemology in the sense of using probability functions to model credal states). Nevertheless, many ways of thinking about the evolution of credence and the connection between justified credence and justified belief will be more naturally compatible with RATIONALISM than DOGMATISM.

2.3 THE BAYESIAN ARGUMENT AGAINST DOGMATISM

Let us consider a formal depiction of a case of dogmatist learning via ampliative inference. Subject S acquires evidence E, has prior background knowledge K, and infers H from E. But, S does not have any justification, prior to learning E, for $E \supset H$. That is, K does not justify belief in $E \supset H$. According to the dogmatist, this is just fine—S need not have any justification for believing $E \supset H$ prior to inferring H from E. Nevertheless, *after* S infers H from E, she will have justification to believe $E \supset H$.

That is, the dogmatist is committed to the following claim, which will be the target of the Bayesian argument.

> TARGET. Possibly, S gains justification to believe $E \supset H$ for the first time when S infers learns E.

Now, to construct the Bayesian argument, we will need to assume PROBABILISM and CONDITIONALIZATION. The following is a provable consequence of the standard probability axioms (see Weatherson (2007) for a proof).

> PROBABILITY THEOREM. If p is a probability function, $p(E \supset H|EK) \leq p(E \supset H|K)$, with equality only if $p(E \supset H|K) = 1$.

This shows that, in the standard lingo, E cannot *confirm* $E \supset H$, and unless $E \supset H$ is already certain, it will actually *disconfirm* it. On a natural understanding of confirmation, this may make it seem like the very same evidence that justifies belief in H cannot possibly justify belief in $E \supset H$, since it should make one less confident in that proposition. But the claim that evidence E confirms a proposition P just in case $p(P|E) > p(P)$ should be taken as a stipulative definition of a technical sense of "confirmation," and it is not immediately clear how this relates to justification to believe. To complete the Bayesian argument, therefore, we need to assume an additional linking principle that connects changes in rational credence to changes in justification to believe, such as the following (modified from Weatherson (2007)).[7]

> Learning. If before learning E S is (1) justified in having credences cr_K and (2) not justified in believing P, and after learning E (and nothing stronger) S is justified in having credences cr_{EK}, then if $cr_{EK}(P) \leq cr_K(P)$, S does not acquire justification to believe P upon learning E.[8]

But if we assume PROBABILISM, CONDITIONALIZATION, and LEARNING, we can conclude that in every case, S does not gain justification to believe $E \supset H$ for the first time when she learns E, for any background knowledge K. So TARGET is false.

Weatherson (2007) proposed a version of LEARNING in order to argue against CONDITIONALIZATION, and suggest a generalization of Bayesianism that might be compatible with dogmatism. However, LEARNING, though intuitive, is not obviously true. In § 2.3.1, I will consider a recent argument from Kung (2010) shows that LEARNING may be false. I do not find his argument perfectly convincing, but I think plausible counterexamples to LEARNING can be constructed along the same lines. However, in § 2.3.2, I will suggest that LEARNING is not a crucial assumption of the Bayesian argument, and propose a revised form of the argument.

2.3.1 KUNG'S BAYESIAN DOGMATISM

Kung (2010, 2) describes his conclusion as follows.

> [I]f you initially have no reason to believe H, then intuitively E can confirm H even though $Pr(H|E) \leq Pr(H)$.

Clearly, Kung is not using "confirmation" in the stipulative way I prefer; what he means is that one might go from a state of having no reason to believe H to having some evidence

[7]Here, and in what follows, I adopt the the following simplifying convention: when I write cr_X, where X is a proposition letter rather than a lower-case subscript indicating a time, I mean the relevant individual's credence function when she has total evidence X. That is, this is still meant to indicate a temporally-indexed function, but saves me the trouble of writing "cr_t where X is S's total evidence at t." I do *not* mean the credence function that S *ought* to have given total evidence X, and I hope no confusion results from this convention.

[8]White's original version of the argument does not dwell on a linking principle, but does include the following claim. "Now if I gain justification for a hypothesis, then my confidence in its truth should increase" (White, 2006, 531).

that tells in favor of H (even if it is not enough to justify belief) while one's credence in H remains unchanged, or even decreases. In the terminology of § 1.2.4, one's evidence for H might become weightier even as the balance of evidence remains the same. The case presented in defense of this intuition is similar to the following (I've simplified it slightly; see Kung (2010, §2) for his full presentation).

> ALIEN CARD GAME. You are observing a strange card game being played by two alien creatures. You know absolutely nothing about the rules of the game, though you are pretty sure it is a game. You also cannot really read the expressions of the strange aliens, or understand their language. You have no idea if the player in front of you is winning, but you can see the strange shapes on the cards he holds. Let W be the proposition that that player will win the hand. Since you have essentially no evidence to support or tell against W (other than the mere fact that the player appears to be playing a game, which presumably is winnable), it is reasonable to have credence $cr_{\emptyset}(W) = \frac{1}{2}$.
>
> Suppose that another observer hands you an odds sheet written in a language you can understand. You match the shapes on the cards in the player's hands to the sheet, and see that O: the odds of that hand winning are exactly 1:1. It is reasonable to have $cr_{\emptyset}(W|O) = \frac{1}{2}$, and so reasonable for your credence to remain the same upon learning this new information.

Kung of course grants that in ALIEN CARD GAME, at no point do you have justification to believe W. And in fact the balance of evidence has not changed, as is reflected in the fact that your credence neither increases nor decreases. But you have gained something epistemically signficant; a *prima facie* reason, in Kung's terminology (as opposed to an "all-things-considered" reason). In terms of credence, you now have weightier evidence for both W and $\neg W$, which is reflected in the fact that your credence in W is more stable.[9]

This is not yet a counterexample to LEARNING, since belief in W is justified neither before nor after learning O. In order to defend DOGMATISM against the Bayesian argument, therefore, Kung proposes the following modification of the case.

> ALIEN CARD GAME II. As in alien card game, but there are four players, so you reasonably have credence $cr_{\emptyset}(W) = \frac{1}{4}$. When you look at the odds sheet, you see that the odds of the hand winning are 1:2. So you reasonably have $cr_{\emptyset}(W|O) = \frac{1}{3}$. So you (reasonably) go from having $cr_{\emptyset}(\neg W) = \frac{3}{4}$ to having $cr_{O}(\neg W) = \frac{2}{3}$.

[9]For example, if another human onlooker who you are sure does not know the rules of the game and hasn't seen the odds sheet tells you, "I bet that alien in front of us is going to win, he looks confident," you might have some reason to believe that he has got a better read on the body language of these aliens than you, and therefore you might reasonably give more credence to W. However, the effect of this evidence on your credence will reasonably be greater if you have not looked at the odds sheet than if you have.

Kung claims that what happens in this case is that "Your credence that an opponent will win decreases from $\frac{3}{4}$ to $\frac{2}{3}$ even though you gain all things considered reason, and perhaps even justification, for thinking that an opponent will win" (Kung, 2010, 12). However, he offers no argument for this claim. Given his definition of "all-things-considered reason", it is plausible if we assume that one has no *prima facie* reason to believe $\neg W$ before looking at the odds sheet, since on his view having an all-things-considered reason seems to just be having a prime facie reason and having justification for a credence greater than $\frac{1}{2}$.[10] But I think there are two reasons to resist the conclusion that this is a counterexample to LEARNING.

First, as is clear, all-things-considered reason to believe P does not necessarily suffice for justified belief. It is, again, true that the evidence in favor of $\neg W$ is weightier after learning what the odds are. But it is not clear to me that you now have justification to believe $\neg W$; the balance of the evidence favors it, but that is not necessarily sufficient for justified belief. In fact, you have gained a reason to be *more* confident that W, though you still should not believe it. It seems to me, in fact, that in this case suspension of judgment is called for both before and after learning O.

Second, it is not clear to me that one has no *prima facie* reason to believe $\neg W$ before looking at the odds sheet. Kung appeals to the Principle of Indifference in both cases to determine rational credences for a state of ignorance. But this is a problem with Kung's approach to constructing counterexamples to LEARNING. For it is not the same to have no posterior evidence for a proposition and to have no *reason* to believe it. If there is a particular prior probability distribution that I *ought* to conform to, then by definition I am justified in having credences that derive from that prior, conditionally on my evidence. So if I have no evidence that confirms some proposition P by the lights of my prior, my credence in P will still be the same as my prior credence, and I will be justified in having that credence. Kung, by contrast, assumes something like the following principle.

> NO REASON. If $cr_\emptyset$ is my prior credence function and if my present evidence includes only evidence that is independent of P (i.e., for all propositions Q that I've learned, $cr_\emptyset(PQ) = cr_\emptyset(P)cr_\emptyset(Q)$), then I am not justified in believing P.

But this is clearly untenable. For one thing, this principle entails that I am never justified in

[10] "When your prima facie reasons to believe P are stronger than your prima facie reasons to disbelieve P, you have an all things considered reason to believe P."

believing any Q with $cr_\emptyset(Q) = 1$. But we certainly may be justified in believing such things (e.g., that $P \supset P$). In fact, presumably we are justified in believing at least some of them in the absence of any posterior evidence. So even if it were possible to build an exception for such beliefs, it will presumably be because I when I *reasonably* have extremely high prior credence in some proposition Q, I may thereby be justified in believing it. There seems to be no obvious reason that "extremely high" here should be limited to propositions that we are certain of, in general.

Kung often seems to suggest that one's prior credences *are not* justified, in the absence of evidence.

> If you initially have no information about H, and hence no reason to believe that H, then when you first gain information you do not *adjust* your credence in H; instead you *replace* your initial credence with one based on the information. It makes sense to adjust your old credences in light of new information when those old credences are based on information that you still regard as worthy of consideration. But if your initial assignment does not represent information that must be weighed and reconciled with the new information, then the initial assignment ought to be discarded as soon as you get *any* information.(Kung, 2010, 5, emphasis in original)

This is suggestive of a view that differs from Kung's, in which, possibly, I should *replace* my priors on learning E in the sense that when I learn E I should adopt $cr_E(H) \neq cr_\emptyset(H|E)$, but instead I should replace my priors with a different function and *then* conditionalize. But in fact, all of his examples are cases of conditionalizing on one's priors rather than adopting new priors, and Kung assumes PROBABILISM and CONDITIONALIZATION throughout; it is the linking principle in the Bayesian argument that is his explicit target.[11] Towards the end of the essay, Kung (2010, 15) explicitly considers several other sensible ways of resisting the Bayesian argument, but does not endorse any of them. We might resist the Bayesian argument by:

- "First, we could suspend judgment under uncertainty, where that means refusing to assign priors when we lack any reasons or information. While I myself am attracted to this alternative, obviously no Bayesian can accept it." Here, "suspension of judgment" clearly does not refer to the ordinary doxastic attitude that goes along with belief and

[11] "The problem with White's argument lies with his justification lemma, that 'if I gain justification for a hypothesis, then my confidence in its truth should increase"' (Kung, 2010, 10).

disbelief. I am not sure how to understand the idea of "refusing" to assign priors. One way might be to have a maximally imprecise, interval-valued credence of $[0, 1]$ in every proposition. See §2.3.4 for more on how this suggestion might be developed.

- Rejecting CONDITIONALIZATION. "The idea would be that conditionalization is appropriate for adjustment, but not replacement." This is the suggestion just considered, and will be explored in more detail in §2.3.3.
- Imposing additional rational constraints on permissible prior credence (beyond the Principle of Indifference and the other usual suspects).

This last possibility is developed in response to a toy example from White (2006). There, White considers a dangerous card game, in which you are put to sleep and one of three cards is drawn at random. Depending on which card is drawn, one of three things will happen. Either (1) nothing will be done to you and you'll be awakened intact, (2) your hands will be cut off and then you'll be woken up, or (3) your hands will be cut off and replaced with superfake hands, which are fake hands that are absolutely indistinguishable from real hands. White uses this to show that, if H is the proposition that you have hands and E is the proposition that you seem to have hands on being awakened, and if you divide your initial credence equally between the three different outcomes (as represented in the figure below), then you'll have credences as in Figure 2 prior to being put to sleep, so that both H and

Figure 2: SKEPTICAL PRIOR PARTITION

EH	$\neg E \neg H$	$E \neg H$

- $cr_{\emptyset}(H) = \frac{2}{3}$
- $cr_{\emptyset}(H|E) = \frac{1}{2}$
- $cr_{\emptyset}(E \supset H) = \frac{2}{3}$
- $cr_{\emptyset}(E \supset H|E) = \frac{1}{2}$

$E \supset H$ will be disconfirmed by E. Kung suggests, by contrast, that DOGMATISM might be saved simply by partitioning the probability space differently: one might permissibly treat H and $\neg H$ as equiprobable, and then further subdivide $\neg H$ into the skeptical and ordinary handless cases. Although the random card element of White's toy case is meant to justify his particular assignment of probabilities, given the principle of indifference, that structure might not reflect the reasonable credence distribution that we should have in these possibilities in the absence of such a contrived setup, after all.[12] So Kung suggests that DOGMATISM might be saved by providing additional constraints on our prior credences, as in Figure 3. Given these priors, H is confirmed by E, and while $E \supset H$ is disconfirmed, it

Figure 3: DOGMATIST PRIOR PARTITION

H	
$\neg E \neg H$	$E \neg H$

- $cr_{\emptyset}(H) = \frac{1}{2}$
- $cr_{\emptyset}(H|E) = \frac{2}{3}$
- $cr_{\emptyset}(E \supset H) = \frac{3}{4}$
- $cr_{\emptyset}(E \supset H|E) = \frac{2}{3}$

is still above $\frac{1}{2}$ and so possibly a candidate for belief.

However, unless LEARNING is false, this is no help to the dogmatist. For then even if we have these priors and are justified in believing both H and $E \supset H$ after learning E, unless the priors themselves are not justified, then LEARNING entails that we'd be justified in believing $E \supset H$ prior to learning E as well.

I find Kung's alien card game cases unconvincing as counterexamples to LEARNING for precisely the reason that in those cases, both before and after getting the relevant evidence it seems to me that one would be justified in having the credences described in the story, but also would not be justified in believing either W or $\neg W$.[13] However, I am still inclined to think that LEARNING is not necessarily true. Here is my counterexample.

[12] It is not clear to me why, once we move outside the card game setup, however, we should not also give some credence to the possibility of having hands and failing to notice them, i.e., $\neg EH$.

[13] I may be being slightly unfair to Kung here. Although I've presented his essay as arguing against LEARNING, in fact his official line in the concluding section is that, while he is sympathetic to some of the three proposals listed above, he is not willing to endorse them. He concludes by endorsing a "fourth option," which I do not agree is distinct from the three above. "There is an intuitive difference between neutrality conferred by balanced evidence and neutrality conferred by lack of evidence. The former calls for adjusting credences, the latter calls for replacing them…Intuitively we want to treat those cases differently. When you replace credences, you can gain all things considered reason, and even justification, despite the fact that your updated credences are less than your prior credences" (Kung, 2010, 15–16). As noted above, the phrase "and even justification" seems to be doing the work of a missing premise. Given the way Kung

RAIN. At time t, I look at the sky and become confident that it is going to rain (R). Suppose that my credence $cr_K(R) = .9$ and that I dislike getting wet and also dislike carrying an umbrella around just enough that my expected value at t for carrying an umbrella is positive just in case $cr_t(R) > .5$. So if I had nothing to go on but my inexpert assessment of weather conditions, I would take an umbrella. But I also know that I've often thought it was likely to rain on days when it did not, and I am not particularly confident in my predictions about the weather. In fact, conditionally on the weather forecast predicting a low chance of rain, I prefer not to take an umbrella. So it is not rational for me to act as though R at t, given that I can very very easily check the chance of rain on the internet before leaving home. So I am not justified in believing R at t. I do check the weather forecast, and I learn at time t' that the chance of rain today is 80%. My credence in R decreases from .9 to .8. But now it is rational for me to act as though R. So I at least satisfy the pragmatic constraint on (justified) belief discussed in §1.2.3, and intuitively, I am justified in believing R at t'.

This example relies on pragmatic constraints to show that I am not justified in believing R at t, so it may not convince anyone who does not think there are pragmatic constraints on (justified) belief.[14] But I think the intuitive judgment may hold up even if you do not buy the pragmatic explanation—at t, when all I have to go by is my inexpert assessment, I am not justified in believing R, whereas at t' I am. I think it is also plausible that I am justified in having the credences I do at both times. There is no harm in relying on one's inexpert assessment of the evidence, particularly when one is forthright about one's fallibility and seeks corroboration.

Nevertheless, I do not think that DOGMATISM can be saved by simply by stipulating that the dogmatist prior partition is permissible, while still holding onto CONDITIONALIZATION and rejecting LEARNING instead. In the next section, I will suggest a replacement principle

treats all-things-considered reasons, he may be right about the conclusion with "justification" left out, but then this is irrelevant to LEARNING. And perhaps this is all he means to defend. "Thus even if White is correct that your credence that you are deceived increases upon having an experience as of hands, the dogmatist can still make sense of the claim that we gain all things considered reason to believe that we are not deceived." However, unless all-thing-considered reason in favor of P confers justification for belief in P, this is an insufficient defence of DOGMATISM.

[14] As I claim in §1.2.3, I do not think that the reduction of belief that P to credence sufficient for invariance in conditional preferences over the things that matter on the assumption that P is successful. Nevertheless, I am open to the view that variance in conditional preferences over the things that matter on the assumption that P is incompatible with justified belief that P. In this case, I am not justified in believing R prior to checking the weather forecast, because conditionally on the weather report saying that the chance of rain is only 40%, I prefer not to take an umbrella, but conditionally on the weather report saying that the chance of rain is only 40% *and* R, I prefer to take one. Once I have seen that the weather report says that the chance of rain is 80%, the possibility that it says the chance is only 40% is no longer salient—it is not one of the things that matter. Both before and after checking the weather, I prefer to take an umbrella given my credence in R, but only after checking the weather report are my conditional preferences invariant on the assumption of R (since after checking the weather, conditionally on all salient possibilities, I prefer to take an umbrella).

for LEARNING that is not vulnerable to these kinds of counterexamples (in part because it applies specifically to belief in conditionals, like $E \supset H$). The revised Bayesian argument will show that the only hope for DOGMATISM is to reject CONDITIONALIZATION, or a probabilistic approach to epistemology altogether.

2.3.2 THE BAYESIAN ARGUMENT REVISED

I agree with Kung (2010) that LEARNING is false, though not for the same reasons. But I've suggested that the dogmatist should receive cold comfort from this fact. For we can replace LEARNING with a different premise that I think is correct. It does not apply as generally as LEARNING, but it will apply specifically to the case of concern for DOGMATISM, i.e., the case of gaining justification for $E \supset H$ when one learns E. First, let me offer a definition.

> CONDITIONAL JUSTIFICATION. For any proposition P, a subject S is *conditionally justified* at t in believing that P given E just in case S is justified in having credences cr_t such that, if S were to have (at t) $cr_t = cr_t(\cdot|E)$, S would believe that P.

Now, if LEARNING is false, then even if we assume PROBABILISM and CONDITIONALIZATION, it is quite possible that S could be conditionally justified at t in believing that P given E even though $cr_t(P|E) < cr_t(P)$. For example, in RAIN, I am conditionally justified in believing that R given that the weather report says that the chance of rain is 80%, but my credence in R conditional on that evidence is lower than my unconditional credence.

My argument will also rely on the following assumption.

> P1. If S is conditionally justified at t in believing that P given A and S is conditionally justified at t in believing that P given B, then S is conditionally justified at t in believing that P given $A \vee B$.[15]

Observe also that the following principle must be true, given my definition of conditional justification.

> PRE-JUSTIFICATION. If S is conditionally justified at t in believing that P given $\top$ (a logical truth), then S is justified in believing that P at t.

[15] I confess I am not completely confident that an intuitive counterexample to this premise cannot be found, though so far I have failed to think of one. Imagine such a case would be one in which A and B are both undercutting defeaters for a justification for P based on the other, so that having $cr_t(\cdot|AB)$ would not justify belief in P. But the cases I care about this cannot happen, since A and B are mutually exclusive, so I expect that my premise is at least true enough.

PRE-JUSTIFICATION must be true, for if S is conditionally justified at t in believing that P given $\top$, then by definition if S were to have $cr_t = cr_t(\cdot|\top)$, S would believe that P. But of course, this identity is necessarily true, so assuming S's credences are justified, then S is justified in believing that P.

With these two premises and PROBABILISM and CONDITIONALIZATION, we can exploit the truth conditions for the material conditional to show that DOGMATISM is false. For if CONDITIONALIZATION is true, then were she to learn E, her credence function should just become $cr_t(\cdot|E)$. The dogmatist says that if S had that credence function, she would be justified in believing that $E \supset H$. But as long as S knows that $E \supset H$ is entailed by $\neg E$, then S is also conditionally justified at t in believing that $E \supset H$ given $\neg E$. For $cr_t(E \supset H|\neg E) = 1$ (assuming PROBABILISM). So by P1, S is conditionally justified at t in believing that $E \supset H$ given $E \vee \neg E$. But then S is justified in believing that $E \supset H$ at t, prior to learning E. Since this must be true in every case in which it is possible to gain justification for $E \supset H$ by learning E, this contradicts DOGMATISM.

Let me offer an example. Consider that if DOGMATISM is correct, and if it is true that the fact that all observed emeralds are green justifies belief in the conclusion that all emeralds are green, then I might reason to myself in the following fashion. If I were to learn that all observed emeralds are green, I would be justified in believing that all emeralds are green. I would also be justified (by JUSTIFICATION CLOSURE) in believing that [if all observed emeralds are green, then all emeralds are green]. But if I were to learn that it is not the case that all observed emeralds are green, it would be vacuously true that [if all observed emeralds are green, then all emeralds are green], so I would be justified in believing it. So no matter what I learn about the color of observed emeralds, I would be justified in believing that [if all observed emeralds are green, then all emeralds are green]. So since $E \supset H$ would be justified for me no matter what evidence I were to get regarding the color of observed emeralds, I ought to just go ahead and believe it now (of course, I will not be justified in believing H without finding out whether E).

Now, I might not be able to *perform* this particular piece of reasoning, because I may not know the crucial premise, "If I were to learn that all observed emeralds are green, I would be justified in believing that all emeralds are green." After all, this is just the sort of thing that inferential externalists, of which dogmatists are a species, do not think we have to know to gain justification. Nevertheless, the soundness of this sort of reasoning, if DOGMATISM is true, suggests that I would have propositional justification for $E \supset H$, given that I am conditionally justified in believing $E \supset H$ given $\top$.

So both LEARNING and P1 entail with PROBABILISM and CONDITIONALIZATION that DOGMATISM is false. Take your pick; I prefer P1 for the reasons given in §2.3.1, but possibly my counterexample to LEARNING is unconvincing, making the retreat to P1 unnecessary.

Given P1, the dogmatist has only two options. First, she might object to my probabilistic approach altogether, and insist that the notion of propositional justification she is interested in is not captured by anything in the Bayesian framework. I will ignore that possibility in favor of (what I think most dogmatists would prefer) trying to find a probabilistic home for DOGMATISM. In particular, the dogmatist might reject CONDITIONALIZATION. For I relied on CONDITIONALIZATION to generate the crucial premise that if S is justified in believing that P after learning E, then S is conditionally justified prior to learning E in believing P given E. In fact, there are many unorthodox Bayesians who are skeptical of CONDITIONALIZATION,[16] and some proponents of dogmatism have already proposed dropping the orthodox Bayesian story about how probabilities should change over time.[17] The next two sections consider two different proposals for doing just this.

2.3.3 REJECTING CONDITIONALIZATION: OBJECTIVE BAYESIANISM

The Bayesian argument shows that, if we assume PROBABILISM and CONDITIONALIZATION, then if S is justified in believing that $E \supset H$ after learning E, S was independently justified in believing $E \supset H$ prior to learning E. This is because the post-facto justification that S acquires whether E or $\neg E$ is conditionally available to S given her prior credences. However, this assumes that those credences are justified; if they are not, the beliefs they give rise to will not be justified either (and simliarly for conditional credences and conditional justification). As noted in §2.3.1, one maneuver available to the dogmatist is to propose that only priors that give high credence to $E \supset H$ are permissible. And Kung's talk of "replacing" credences

[16] Cf. Elga (2000); Arntzenius (2003); Christensen (2007a); Weisberg (2009).
[17] Cf. Pryor (unpublished manuscript), Weatherson (2007).

rather than "adjusting" them upon learning E similarly suggests that when one's priors are *not* justified, one should not obey CONDITIONALIZATION, but should instead replace one's credence function with a new, justified credence function. That is, his second suggestion (and perhaps the third as well) can be understood as advocating objective Bayesianism, which replaces CONDITIONALIZATION with a less subjective version.

> OBJECTIVE CONDITIONALIZATION (OC). For all propositions P, and subjects S, S ought at t to have $cr_t(P) = o(P|E)$, where o is a permissible prior probability distribution, and E is S's total evidence at t.

For simplicity, let me assume in this section that there is a *unique* permissible prior probability distribution o, which I will sometimes call the "One True Priors." On the proposal in question, it is possible for S to go from having an unjustified credence distribution at t to having a justified credence distribution that permits belief in both H and $E \supset H$ when S learns E. For suppose S has $cr_K(E \supset H) < \frac{1}{2}$, so that S does not believe $E \supset H$. By PROBABILITY THEOREM, if S learns E and obeys CONDITIONALIZATION, S will not believe $E \supset H$ after learning E, either. But suppose that the One True Priors are such that $o(E \supset H|K) > \frac{1}{2}$ and $o(E \supset H|EK) > \frac{1}{2}$. Then by OC, cr_K is not justified for S. But if S (correctly, by OC) does *not* conditionalize on her own priors when she learnes E and instead adopts $cr_{EK} = o(\cdot|EK)$, then S's credences will be justified after learning E, and, possibly, S will be justified in believing $E \supset H$.

This may be what the liberal about inferential justification wants. However, on this view, S had *propositional* justification to believe $E \supset H$ even prior to learning E. Recall JC from § 1.2.1.

> JC. $cr_t(P) = x$ is *justified* for a subject S at time t if and only if $\exists p$ with $p(P|E) = x$, where E is S's total evidence at time t and p is a permissible prior probability function.

Let cr_K rigidly refer to S's actual credences at t given her background knowledge K. Then by hypothesis, in the dogmatist case described above, $cr_K \neq o(\cdot|K)$. This is just what it means for a credence distribution to be unjustified for S at t, on the objective Bayesian assumption that o represents the One True Priors. But JC also entails that $o(\cdot|K)$ is justified for S at t. So S epistemically ought to give high credence to $E \supset H$, whether she does or not. This is not quite yet to say that $E \supset H$ is propositionally justified for S at t. But now on the assumption that having $cr_t(E \supset H|E) = o(E \supset H|EK)$ is epistemically required of S at t, then S is conditionally justified at t in believing $E \supset H$ given E. But S also should have $cr_t(E \supset H|\neg E) = 1$ (this is, again, true of *any* probability function, including o). And now we can run the (revised) Bayesian argument again, to show that S is justified in believing $E \supset H$ at t, because S is conditionally justified at t in believing $E \supset H$ given $E \vee \neg E$.

So although the suggestion that we should sometimes replace impermissible credences with justified ones does allow for the possibility of, for the first time, having a justified belief that $E \supset H$ when one learns E, it entails that we nevertheless have justification to believe $E \supset H$ independently of, and prior to, learning E. We may still fail to have a well-founded belief that $E \supset H$, or even an *ill*-founded belief. We may, therefore, not have a belief that $E \supset H$ available to employ in inferring H from E. So this suggestion may still be compatible with LIBERALISM, in that a *prior* justification to believe $E \supset H$ may play no role in justifying the transition from $cr_K \neq o(\cdot|K)$ to $cr_{EK} = o(\cdot|EK)$.[18] Nevertheless, on this objective Bayesian story, we do have prior and independent propositional justification to believe $E \supset H$ whenever it is possible to infer H from E and thereby gain a justified belief that H. So this story may be compatible with a rationalist LIBERALISM, but it is not compatible with DOGMATISM. Rather, the view seems to be that we *always* have a presumptive justification to believe whatever we would believe if our priors were the One True Priors.

2.3.4 REJECTING CONDITIONALIZATION: DYNAMIC KEYNESIANISM

The last of the suggestions for saving DOGMATISM from Kung (2010) is that we might be justified in "suspending judgment" under uncertainty, i.e., "refusing to assign priors" in the absence of information that determines what they ought to be. I have some sympathy for this proposal, if it is understood as the claim that when I have no information that constrains my rational prior credence in P, I ought to have a maximally imprecise credence $cr_{\emptyset}(P) = [0, 1]$.

It is already common to generalize the Bayesian model of credal states that assign to every proposition a single sharp-valued probability in order to accommodate imprecise or interval-

[18]Even this is not clearly a possibility; it depends on how we describe "what justifies" this credal replacement. This ambiguity about what "does the justifying" is already present in LIBERALISM, however.

valued credences. The motivations for doing so are various. Some (e.g., Hájek (1998)) argue that it is simply more psychologically realistic not to require sharp credences; sometimes, for example, we may not even be able to assess whether we think P more likely than Q, let alone assign a perfectly precise numerical value to our probability assessments. Some (e.g., Weatherson (2007)) suggest that this allows us to capture the Keynesian distinction between risk and uncertainty. The best argument, in my opinion, is presented by Joyce (2005) and Sturgeon (2008). They argue that we need imprecise credences because sometimes the evidence does not determine any precise sharp-valued credence, and so only an imprecise credence fits the evidence.[19]

The typical way of modeling imprecise credences is with a set of probability functions called the subject's "representor." So, without violating its spirit or abandoning a generally probability-first approach, we replace PROBABILISM with

> KEYNES. The doxastic state of an ideally rational subject at any time is represented by a set of probability functions.

If $\mathcal{R}$ is S's representor, then one can say that S's credence in a proposition P is $\{x : \exists p \in \mathcal{R}$ with $p(P) = x\}$. The typical view does not, however, abandon conditionalization. On what Weatherson (2007) calls the "static Keynesian" model, what one does when one gains evidence E is simply conditionalize every function in $\mathcal{R}$ on E.

Note that this sort of model might in fact allow us to explain the failure of LEARNING in my RAIN case, as well. For let P be the proposition that it will rain, and E be the proposition that the weather report says the chance of rain is 80%. I claimed previously that I might reasonably not believe P despite having $cr_t(P) = .9$, because I lack confidence in my inexpert assessment of the evidence. But perhaps this sort of uncertainty is better modeled by an interval-valued credence; perhaps my credence in P should be $cr_t(P) = (.4, .9)$, say. While the precise relationship between credence and belief is no clearer on an imprecise model than it is on a sharp model, just as a sharp credence of $\frac{1}{2}$ plausibly mandates suspension of judgment, presumably an imprecise credence that includes values both below and above

[19] Elga (2010) and White (2010) argue that credences should be sharp, against this orthodoxy. Elsewhere I have argued in favor of the imprecise model, on the grounds that the particular sharp credences that we ought to have if credences should be sharp are sometimes *incompatible* with the evidence.

$\frac{1}{2}$ also plausibly mandates suspension of judgment.[20] However, it will also be true that $\forall p \in \mathcal{R},\ p(P|E) = .8$ (assuming I take the weather report to provide an accurate assessment of the chance of rain). So on learning E, my imprecise credence in P collapses into a precise credence of .8. And so perhaps this explains why I am justified in believing P after learning E, but not before. So while LEARNING may be false if credences should not be sharp, perhaps an analogous version for a Keynesian model *is* true. Something like:

> LEARNING*. If before learning E S is (1) justified in having $cr_K(P) = (x, y)$ and (2) not justified in believing P, and after learning E (and nothing stronger) S is justified in having $cr_{EK}(P) = (w, z)$, then if $x \leq w$ and $z \leq y$, S does not acquire justification to believe P upon learning E.

The revised RAIN case above is compatible with LEARNING*. But since $p(E \supset H|E) \leq p(E \supset H)\ \forall p \in \mathcal{R}$, by PROBABILITY THEOREM, then LEARNING* entails that DOGMATISM is false if we assume CONDITIONALIZATION and KEYNES.

Weatherson's purpose in presenting the Bayesian argument is not to argue against DOGMATISM, but rather to motivate a further generalization of the Keynesian model by allowing for a different sort of operation that can run in parallel with conditionalization. On this model, when one gains evidence E, one may appropriately respond to this evidence by throwing out some of the functions in one's representor.[21] One then conditionalizes all the functions that remain in one's representor on the new evidence.

This allows for the following sort of case. Suppose S has $cr_t(E \supset H) = (.4, .9)$, with $p \in \mathcal{R}$ such that $p(E \supset H) = .85$ and $p(E \supset H|E) = .8$. Suppose also that $U(\mathcal{R}, E) = \{p\}$. Then on getting evidence E, although p assigns a lower credence to $E \supset H$, S now has a high sharp-valued credence in $E \supset H$ where S previously had an imprecise credence that spanned even lower values. Thus it is not clear that we should say that learning E always lowers credence in $E \supset H$, even in the LEARNING* sense, on the dynamic Keynesian model.

[20] As I briefly mention in §1.2.4, the third property of evidence that Joyce (2005) discusses is its *specificity*. The idea is that more specific evidence requires sharper credence. Highly unspecific evidence will typically require suspension of judgment, but evidence need not be perfectly specific to be decisive; possibly an interval-valued credence of, say, (.9,.99) might be compatible with belief.

[21] Formally, one's representor becomes a new set $U(\mathcal{R}, E)$, where U is a function that takes as inputs sets of probability functions and propositional evidence and returns a new set of probability functions that is a subset of the input set. It is not entirely clear why Weatherson requires the output to be a subset of the input—while this is in the spirit of Bayesianism, in fact it seems possible that new evidence might *increase* uncertainty.

Weatherson offers the following proposal for when we should say that one's credence in a proposition is lowered. A subject S with representor $\mathcal{R}$ has her credence in P *lowered* on learning E if:

- For all cr in $U(\mathcal{R}, E)$, $cr(P|E) < cr(P)$.
- For all cr in $\mathcal{R}$ but not in $U(\mathcal{R}, E)$, there is a cr' in $U(\mathcal{R}, E)$ such that $cr'(P|E) < cr(P)$.

In the toy example above, there must have been a function $q \in \mathcal{R}$ with $q(E \supset H) = .5$, but no function in $U(\mathcal{R}, E)$ assigns a lower value to $E \supset H$ conditional on E than q did unconditionally, so the second criterion is not fulfilled, and S's credence in $E \supset H$ is not lowered. In fact, in the toy case above, S's credence in P undergoes precisely the same dynamic I suggested it might in my revised RAIN case. So dynamic Keynesianism may not be vulnerable to the Bayesian argument (or its Keynesian counterpart using LEARNING*).

Can we fall back on an analog of the revised Bayesian argument here? If p represented S's credences all along, S would be conditionally justified in believing $E \supset H$ given $E \vee \neg E$. But we cannot say that the other functions in $\mathcal{R}$ would have the same feature; it is because they have been thrown out by the revision function U that S is justified in believing $E \supset H$ after learning E, on the dynamic Keynesian story. So it may be that prior to the operation of this revision function U, S is *not* conditionally justified in believing $E \supset H$ given E.

Dynamic Keynesianism is the best hope for a probabilistic version of DOGMATISM. Weatherson describes the virtues of his model thus:

> Intuitively, [the revision function U] models the effect of learning, via getting evidence E, what evidential relationships obtain. In the static Keynesian model, it is assumed that before the agent receives evidence E, she could already say which propositions would receive probabilistic support from E. All of the relations of evidential support were encoded in her conditional probabilities. There is no place in the model for learning about fundamental evidential relationships. In the dynamic Keynesian model, this is possible. When the agent receives evidence E, she might learn that certain functions that were previously in her representor misrepresented the relationship between evidence and hypotheses, particular between evidence E and other hypotheses. In those cases $U(R, E)$ will be her old representor R, minus the functions that E teaches her misrepresent these evidential relationships. (Weatherson, 2007, 178)

What is not clear, though, is how learning just E could be instructive about these evidential relationships. How does learning E justify us in discarding functions from our representor that give low prior probability to $E \supset H$? Here is one suggestion: only some prior probability distributions are permissible, and at all times t, we ought to include in our representor only those credence functions that are derived from a permissible prior via conditionalization. So it is always rational to discard other functions that misrepresent evidential relationships because they are derived from impermissible priors. This would be to combine Keynesianism with OC, however, and would be inconsistent with DOGMATISM for the same reason that Bayesian OBJECTIVE CONDITIONALIZATION was: even if we have not yet thrown out the impermissible functions, we ought to, and we always have a presumptive propositional justification to believe $E \supset H$ since every permissible prior is such that we are conditionally justified in believing $E \supset H$ given $E \vee \neg E$. E plays no special role. If DOGMATISM is to be viable, the revision function cannot operate simply by weeding out credence functions with priors that were *all along* impermissible. It has to be that somehow, learning E makes it *now* rational to discard some functions when it was not before.

But why would learning E make it rational not just to conditionalize on E but also, say, to discard functions that have $p(H|E) = p(H)$, keeping those that have $p(H|E) > p(H)$? To take a more concrete example, suppose among the functions in my representor are p and q. Let H be the hypothesis that all emeralds are green, and H' be the hypothesis that all emeralds are grue. Let E be the proposition that all observed emeralds are green. Suppose $p(H) = \frac{1}{2} = q(H')$, $p(H|E) = \frac{9}{10} = q(H'|E)$. That is, by the lights of p E confirms H, while by the lights of q the same evidence confirms H' to exactly the same degree. The sort of thing that the revision function U is supposed to do is make it rational to discard functions like q, i.e., $q \notin U(\mathcal{R}, E)$.[22] For if I have both of these functions in my representor and learn E and *merely* conditionalize, I will not wind up believing that all emeralds are green (my imprecise credence in this hypothesis will include values both higher and lower than $\frac{1}{2}$). But why would learning E *and nothing else* make it rational to throw q out of my representor, on the assumption that I reasonably had q in my representor before learning E? If I begin

[22]See, e.g., Weatherson (2007, 182). There Weatherson describes the mental life of very strange disembodied souls, and suggests that there is no reason we should know a priori what their evidence (from empathic connections with other souls) epistemically supports. "Note that to answer this question, we'd have to know which of these concepts were grue-like and which were projectible, and there is no reason to believe we are in a position to know that." So when the revision function U "shrink[s] our representors to sets of probability functions which are broadly speaking epistemically appropriate for the kind of world we are in," one of the things it presumably does is toss out the functions on which the evidence supports hypotheses involving grue-like predicates.

by being reasonably uncertain whether E is evidence for H or H', and then I learn E, it strikes me as epistemically impermissible to subsequently believe H (and that E is evidence for H and not for H'), assuming that I learn nothing stronger than E.

Let me try to make this more precise. We can define a notion of conditional justification for the Keynesian generalization of Bayesian epistemology as well.

> Keynesian Conditional Justification. For any proposition P, a subject S is *conditionally justified* at t in believing that P given E just in case S is justified in having a representor $\mathcal{R}$ such that, were S to have (at t) $\mathcal{R}' = \{p' : p' = p(\cdot|E),\ p \in \mathcal{R}\}$, S would believe that P.

This does not appeal to the effects of the revision function U, because being conditionally justified at t in believing that P given E is not the same as being justified in believing that P when one learns that E, but is instead supposed to reflect what, given one's state at t, would be justified conditionally on E. By hypothesis, the effects of E on one's representor that are reflected in the revision function U *are not* part of one's credal state prior to learning that E, and may be unforeseeable.[23] There are some kinds of evidence that we know would remove uncertainty—for example, in the revised RAIN case, I can foresee that learning what the weather report says will remove my uncertainty about R and provide me with a precise assessment of the risk of R, because all permissible prior credence distributions have $cr(R|\langle ch(R) = .8\rangle) = .8$. But in the dynamic Keynesian proposal, there are also kinds of evidence that will remove uncertainty, but where this is not foreseeable. So I might well reasonably take myself *not* to be conditionally justified in believing that P given E, even though were I to learn E, I would be justified in believing that P.

This position has some counterintuitive consequences. For one thing, it implies that we sometimes cannot reasonably direct our own inquiries. For it implies that I might well be justified in believing that conditionally on E, I am not justified in believing that P (or $\neg P$). But then surely it would be foolish for me to go looking for evidence E if I wanted to figure out whether P. And yet, by the dynamic Keynesian dogmatist's lights, that is just what would be successful.

[23] "The whole point of the model is that we can only learn which hypotheses are supported by certain evidence by actually getting that evidence. If we could say just what U is, we would be able to know what was supported by any kind of evidence without getting that evidence" (Weatherson, 2007, 181).

I expect that Weatherson or other defenders of dogmatism will be willing to embrace some counterintuitive consequences, in order to reap anti-skeptical benefits. Moreover, one might try to argue that this consequence is not so counterintuitive. Especially in cases of perceptual evidence, it may be true that we do not have any conception of what would sort of experience would confirm some H before we experience it—particularly in a hypothetical prior that has learned from *no* experience. But this is not quite the case I've described. For the dynamic Keynesian story entails that I might well have a determinate (though imprecise) conditional credence $cr_{\emptyset}(H|E)$. And I might have such a credence that includes $\frac{1}{2}$, so that, conditionally on E, I ought to suspend judgment in H. Moreover, I might be aware of all of these facts about my state of mind. And yet, the dynamic Keynesian dogmatist should say that I might nevertheless gain justification to believe H when I actually learn E (and nothing stronger). So we can actually say something stronger than *just* that the effects of the revision function U may not be foreseeable, which is not by itself so counterintuitive. The problem is that an ideally rational inquirer[24] with perfect access to her own credences who expects to conditionalize could reasonably regard some evidence as irrelevant or not specific enough to settle the question whether P, when in fact this might be just the evidence that would be decisive. And although the proposal under consideration is one that rejects the rationality of conditionalizing under all circumstances, conditional credence nevertheless still has a clear connection to rationality; when $U(\mathcal{R}, E) = \mathcal{R}$, at least, an ideally rational Keynesian inquirer should conditionalize. The dynamic Keynesian model is a generalization of the static Keynesian model, which is itself a generalization of the Bayesian model. It supplements, but does not abandon, CONDITIONALIZATION.

Despite this, I agree that there is no *formal* obstacle to modeling DOGMATISM using a dynamic Keynesian model. But we also can separate two distinct benefits of adopting such a model. First, Weatherson claims it is a virtue of the dynamic Keynesian model that it, unlike a Bayesian or static Keynesian model, allows for learning about evidential relationships. Second, the dynamic Keynesian model is compatible with DOGMATISM. The first of these motivations can be satisfied without embracing the counterintuitive consequences that

[24] I take it that Weatherson's view is that ideal rationality is compatible with, and may even require, imprecise credence prior to the operation of the revision function U.

come along with a specifically *dogmatist* interpretation. In fact, it can be satisfied without using a dynamic Keynesian model at all; Weatherson's claim that Bayesianism and static Keynesianism cannot accommodate learning about evidential relationships is false.

Consider the following alternative story. Suppose that there is a function that describes the One True Priors, o. Surely o has the following feature: $\forall A \forall B$, $o(A|B\langle o(A|B) = x\rangle) = x$.[25] If we should obey OBJECTIVE CONDITIONALIZATION, then, it follows that we should have subjective priors such that $cr_{\emptyset}(A|B\langle o(A|B) = x\rangle) = x$. But then, if we were to learn that $o(A|B) = x$, this would transform our understanding of the relationship between A and B, if we did not already have $cr_{\emptyset}(A|B) = o(A|B)$. So learning about the One True Priors would be a way of learning about evidential relationships; the One True Priors just reflect the correct take on what is evidence for what. What might cause us to become certain about what the One True Priors are is another question,[26] but this is no more mysterious than how it is that $U(\mathcal{R}, E)$ is supposed to make it reasonable to throw out grue-confirming functions. One possibility is that "seeing" that E makes probable H as per Fumerton (1995) just gives us insight into the One True Priors.

This alternative story is not compatible with DOGMATISM. For, by relying on OBJECTIVE CONDITIONALIZATION, it once again supposes that we have prior justification for $cr_{\emptyset} = o$, and so quite possibly prior justification for $E \supset H$ when E is evidence for H. Note, however, that we might not be justified in believing $\langle o(H|E) = x\rangle$ and may acquire propositional justification to believe it that we previously lacked. For there is no reason to suppose that the One True Priors are certain about what they are: possibly $o(\langle o(H|E) = x\rangle) \neq 1$. So this shows that it is possible to incorporate learning about fundamental evidential relationships into a Bayesian model.

[25] Suppose $\exists A \exists B\ o(A|B\langle o(A|B) = x\rangle) = y, y \neq x$, and suppose that $o(A|B) = x$. Let C be the conjunction of B and $\langle o(A|B) = x\rangle$, so that $o(A|C) = y$. Then assuming OC, if S's total evidence at t is B, S should have $cr_t(A|B) = x$. If between t and t' S learned that the only permissible prior probability function o is such that $o(A|B) = x$, and nothing stronger, so that S's total evidence were to become C, then by OC S should have $cr_{t'}(A|C) = y, y \neq x$. That would be to say that learning *only* that the unique permissible prior probability function is such that the probability of A on your evidence is x should cause you to adopt a credence in A different than x. This is absurd.

[26] Another way of putting this: if there really is a unique permissible prior probability function, then presumably none of us has permissible priors. We'd surely like to have them, but on this demanding conception of rationality, it is not clear what we who are ignorant about the One True Priors should try to do if we want to do better.

So DOGMATISM is compatible with Weatherson's dynamic Keynesian model, and incompatible with the other probabilistic models of updating that we have considered. However, the dynamic Keynesian model has some counterintuitive consequences, and one of the main virtues it has is by no means unique to it—we can model learning about evidential relationships with an objective Bayesian model. In fact, I will show in the next section that we can model learning about evidential relationships *without* abandoning CONDITIONALIZATION for OC.

2.4 CONDITIONALIZATION AND LIBERALISM

So far, I've argued that anti-skepticism about inferential justification requires us to believe that we have justification to believe $E \supset H$ whenever we have justification to to believe H on the basis of an inference from evidence E against background knowledge K. The dogmatist claims that this justification is itself dependent on E. I have argued that this is incompatible with formal Bayesian models of inferential learning. While a dogmatist might adopt the dynamic Keynesian model presented in Weatherson (2007), there seems to be little motivation for doing so apart from compatibility with DOGMATISM. If we drop the dogmatist assumption that our justification for $E \supset H$ is dependent on E, then we can give an objective Bayesian account of what justifies belief in this conditional. In this section, I want to present another formal story of how an attitude of belief in $E \supset H$ might be justified, which is not compatible with DOGMATISM but is compatible with either inferential internalism or a rationalist LIBERALISM.

What I have in mind is the generalization of CONDITIONALIZATION proposed by Jeffrey (1965). In the typical Bayesian story, whenever E is a proposition that is part of my evidence, I am justified in becoming certain of it—for $cr_t(E|E) = 1 \; \forall E$. The motivation for Jeffrey's generalization of CONDITIONALIZATION is that sometimes I may acquire evidence E without being certain that E. For example, I may be directly justified, by visual inspection, in believing to degree .6 that this jelly bean is orange when the lighting conditions are poor and my ability to visually distinguish orange from red is compromised, without becoming certain either that the jelly bean is orange (call this O), or of any other proposition P such that $cr_t(O|P) = .6$.[27]

> Jeffrey Conditionalization. If cr_t is a subject's credence function at time t and between t and t' she becomes directly justified in having $cr_{t'}(E) = x$ (and is not directly justified in adjusting her credence in any stronger proposition), her credence in an arbitrary proposition P should become $cr_{t'}(P) = cr_t(P|E)cr_{t'}(E) + cr_t(P|\neg E)cr_{t'}(\neg E)$.

Since the value of $cr_{t'}(E)$ cannot be determined from Jeffrey conditionalization, because it appears on the right-hand-side of the equation, this means that Jeffrey conditionalization, like objective conditionalization, introduces an extra parameter by which the evolution of our credences is constrained—a subject's subjective priors and the propositions that constitute her evidence do not uniquely determine a rational posterior credence. We also need to know the degree to which she is justified in believing the evidence propositions E.

Now imagine a subject who has subjective priors such that she does not take E to be evidence for either H or $\neg H$, and is presently indifferent to these hypotheses, i.e., she has $cr_t(H|E) = cr_t(H) = cr_t(\neg H) = cr_t(\neg H|E)$. How might she learn about evidential relationships on this picture? Suppose that between t and t' S acquires direct justification for $E \supset H$, such that she rationally increases her posterior credence $cr_{t'}(E \supset H)$, but not to certainty. Suppose also that she gains no independent reason to think that E is more or less likely. Then her credence in H conditional on E must go up, since $cr_t(H|E) = 1 - \frac{1-cr_t(E\supset H)}{cr_t(E)}$ (for $cr_t(E) > 0$). So increasing her credence in $E \supset H$ is equivalent to changing her mind about the relevance of E to H.

Thus, all we need to model learning about evidential relationships in a Bayesian framework is to posit that we can gain direct justification for increased credence in $E \supset H$. Just as directly increasing one's conditional credence in H given E also confirms $E \supset H$ (holding credence in E fixed), increasing one's unconditional credence in $E \supset H$ increases the degree

[27] A Bayesian who prefers conditionalization to Jeffrey conditionalization may well object that what happens in such a case is that you become certain of S = "it seems to me that the jelly bean is orange" and R = "lighting conditions are such that my judgments about when something is orange are 60% reliable," so that $cr_t(O|RS) = .6$. I do not claim that conditionalization cannot handle the cases that motivate Jeffrey conditionalization, but it does seem like Jeffrey conditionalization might be a more realistic story about such cases. It is not particularly plausible that I will be certain of R, or that I need to be in order for my perceptual experience to justify some increased degree of credence in O.

to which E confirms H.

Moreover, if we assume JEFFREY CONDITIONALIZATION rather than OBJECTIVE CONDITIONALIZATION, we can be much more permissive about a subject's subjective priors. Perhaps the permissible priors are constrained only by the Principal Principle, the Principle of Indifference, and the other usual suspects. But there is still room for rational constraints on the posterior credences of a well-informed ideally rational subject. If one can somehow get direct justification for increased *posterior* credence in $E \supset H$, that will affect the degree to which E confirms H by the lights of one's present credences, regardless of one's actual temporally prior credences. That is, a view that relies on JEFFREY CONDITIONALIZATION need not suppose that we are *always* justified in giving high credence to $E \supset H$. Possibly some experience could give one direct justification for increased credence in $E \supset H$.[28]

This sort of view would obviously not be a dogmatist view, however. For it posits directly a source of justification for $E \supset H$ which is independent of E. As usual, whether it is compatible with LIBERALISM depends on what it means for the basis of one's justification to believe H to "include" one's justification for another proposition. But we can at least say the following. On this view, it is learning E that affects one's credence in H. Gaining justification for increased credence in $E \supset H$ may reasonably affect one's credence in H not at all; it is just a necessary condition for E to confirm H that the conditional credence of H given E be high. Moreover, one might *already* permissibly have a high credence in $E \supset H$, in which case learning $E \supset H$ in this fashion would not be required. So while a high credence in $E \supset H$ is a formal requirement for a high conditional credence in H given E, it may well not be a precondition on learning H on the basis of E that one ever learn or "see" that there is a connection between E and H (even if we suppose that is possible). So simply positing the possibility of discovering this connection through gaining direct justification for increased credence in $E \supset H$ need not commit us to the denial of LIBERALISM.

If we adopt a probability first approach to epistemology, we need to choose between norms on the evolution of credence such as CONDITIONALIZATION, OBJECTIVE CONDITIONALIZATION, and JEFFREY CONDITIONALIZATION. None of these is compatible with DOGMATISM.

[28]According to the account of inference to the best explanation in the next chapter, for example, perhaps one could be so justified by having the experience of finding H to be a satisfying explanation of E.

The only model we have surveyed that is compatible with DOGMATISM is the dynamic Keynesian model. This model has the counterintuitive consequence that we sometimes cannot appropriately direct our inquiry towards relevant evidence, because we will expect that the very evidence that would justify belief in H would not do so. Moreover, it is poorly motivated independently of a desire for consistency with DOGMATISM. If a probability first approach to epistemology is correct, then RATIONALISM is more plausible.

3.0 INFERENCE TO THE BEST EXPLANATION

3.1 INTRODUCTION

Just as Bayesianism has been thought to conflict with the dogmatist story about how our beliefs are justified, it is been thought to conflict with inference to the best explanation (IBE), in the sense that a rational subject who obeys PROBABILISM and CONDITIONALIZATION has no room to *also* infer the best explanation of her evidence, if that is to be understood as something other than simply conditionalizing. This thesis has been most famously defended by van Fraassen (1989, chap. 7). To assess this claim, it is important to make several distinctions.

First, by inference to the best explantion, I do not necessarily mean a process that results in *belief* in the best explanation. The most basic objection to IBE is that the best explanation (that we can think of) need not be very good, and in particular not good enough to inspire much confidence. So we often should *not* infer the best explanation, if that means that we should believe it. It is now common to avoid this objection by shifting to a probabilistic framework. We can then recast IBE as a rule of inference (if that is what it is) that obliges us to regard an explanatory hypothesis H as likelier, when we learn E, than any alternative worse explanatory hypothesis H'.

To make clearer exactly what IBE is supposed to be, we must distinguish between the *likeliest* explanation of some evidence, the *loveliest* explanation, and the most *virtuous* explanation of some evidence. I will use these terms in a stipulative way that may differ from some of the literature I will consider.

LIKELINESS. An explantory hypothesis H is subjectively *likelier* than an alternative H' relative to some evidence E, for a subject S with background knowledge K, if and only if

$cr_K(H|E) > cr_K(H'|E)$.

LOVELINESS. An explanatory hypothesis H is *lovelier* than an alternative H' relative to some evidence E, for a subject S, if and only if H strikes S as a better explanation than H'.[1]

VIRTUE. An explanatory hypothesis H is *more virtuous* than an alternative H' relative to some evidence E if and only if H scores better with respect to the explanatory virtues than H'. The explanatory virtues will be traits such as simplicity, unification, elegance, etc.

I know of no adequate account of the explanatory virtues, but proponents of IBE tend to think there are such things and that in principle one ought to be able to describe what makes an explanation better than another. I am somewhat skeptical. I think that some explanations are clearly *lovelier* than others, but it is not clear exactly what we are responding to when an explanation strikes us as lovely. As we will see, one possibility is just that the lovelier explanation is lovelier because it is *likelier*. Nevertheless, inference to the best explanation can be understood as any of: inference to the likeliest explanation, inference to the loveliest explanation, or inference to the most virtuous explanation.

INFERENCE TO THE BEST EXPLANATION. When H is the likeliest/loveliest/most virtuous explanation of some evidence E, for a subject S with background knowledge K, if S learns E (and nothing stronger) S ought to have $cr_{EK}(H) > cr_{EK}(H')$ for all alternative explanatory hypotheses H'.[2]

The heart of the alleged conflict between IBE and Bayesianism has to do with which of these interpretations is correct. If we should conditionalize, then so long as S has probabilistically coherent credences cr_K,[3], her new credences on learning E should be $cr_{EK} = cr_K(\cdot|E)$. Thus, if $cr_K(H|E) \leq cr_K(H'|E)$, then by CONDITIONALIZATION, when S learns that E, she should adjust her credences in such a way that $cr_{EK}(H) \leq cr_{EK}(H')$. So if CONDITIONALIZATION is true, then IBE is a correct norm on inference only if it is understood as inference to the likeliest explanation; if we interpret IBE as inference to the loveliest or most virtuous

[1]The main purpose of this definition is to distinguish between explanations that *seem* good and explanations that really *are* good.

[2]There is admittedly some vagueness in what counts as an alternative explanatory hypothesis relative to H. For instance, does $\neg H$ count? Is the "null hypothesis" an alternative? These issues will arise at some points in the arguments below, where I hope to deal with them as needed.

[3]Recall that I use this alternative subscript notation as follows: cr_X denotes the credences of a subject at the time that her total evidence is X. Thus, if t is the time just prior to learning E and t' is the time just after, and S learns nothing else in between t and t', then $cr_K = cr_t$ and $cr_{t'} = cr_{EK}$.

explanation, then it will be correct only when those properties line up with likeliness. We will then be able to say that the reason we should infer the best explanation is just that we should conditionalize, and the best explanation is the one that is likeliest on the evidence, given our priors. Either inference to the loveliest explanation just is always CONDITIONALIZATION, or sometimes we should not do it. Bayesianism gives the complete story, and IBE is, at best, a poor way of telling it, and, at worst, dangerously misleading.

Friends of IBE, however, have typically wanted to defend the claim that explanatory virtue is a mark of truth, and that explanatory loveliness is a sign of virtue—that we can find reason to prefer hypothesis H to hypothesis H' in the fact that H strikes us as a better explanation, even if we cannot say precisely what makes it lovelier.

The most popular form of defense of IBE accepts that inference to the likeliest explanation is always correct, but seeks to preserve an important role for explanatory loveliness and virtue in our actual practice of inquiry. Lipton (2004); Okasha (2000); McGrew (2003) have argued that IBE is a useful heuristic by means of which we approximate the rational ideals of Bayesianism. This sort of compatibilist view accepts CONDITIONALIZATION, but tries to argue that this does not rob IBE of all interest. As we will see in §3.2.1, the proponents of the heuristic view envision two roles for IBE. First, when $cr_K(H|E)$ is well defined, the fact that H is a good explanation of E somehow determines or reveals that it has a relatively high value. Second, when $cr_K(H|E)$ is not well defined, the fact that H is a good explanation of E either justifies or causes the adoption of a high credence in H conditional on E. I will argue that this attempt at compatibilism fails, and any appearance to the contrary can be explained by a failure to distinguish between different interpretations of the relevant probability functions.

Like dogmatists, friends of IBE may also choose to reject the orthodox Bayesian picture in which CONDITIONALIZATION provides the sole norm on the evolution of credence in response to the evidence. Weisberg (2009) offers a version of this defense of IBE that rejects CONDITIONALIZATION in favor of OBJECTIVE CONDITIONALIZATION, and by extension rejects inference to the *likeliest* explanation.[4] On this view, there are times when S has

[4] Although it will sometimes not be the most natural reading, for the sake of consistency it is important to always read "likeliest" as "likeliest by the lights of one's subjective priors"—a failure to do this is, in my view, the primary source of apparent plausibility for the heuristic view.

$cr_K(H|E) \leq cr_K(H'|E)$ for some alternative H', but she should abandon conditionalization and have $cr_{EK}(H) \neq cr_K(H|E)$. That is, the goodness of H as an explanation of E can justify a high credence in H even though one previously did not have a high conditional credence in H given E.

This section will explore these three different responses to the alleged possibility of conflict between CONDITIONALIZATION and IBE. I agree with Weisberg that IBE does conflict with CONDITIONALIZATION and ought to be understood as a source of rational constraint on the evolution of credence that is distinct from CONDITIONALIZATION. I will then develop an account of how this constraint operates that is compatible with a rationalist liberal view about inferential justification, and defend it from some skeptical worries. In particular, I will claim that we are sometimes immediately justified in believing $E \supset H$ because H is an intellectually satisfying explanation of E.

3.2 VAN FRAASSEN'S ARGUMENT AGAINST INFERENCE TO THE BEST EXPLANATION

The argument against IBE in van Fraassen (1989) relies on a typical Dutch Book to support CONDITIONALIZATION. Van Fraassen shows that a person who does not update her credences via conditionalization is vulnerable to a set of bets that she would regard as fair given her credences, but which would lead to a guaranteed loss. As with all Dutch Book arguments, this is supposed to be a demonstration of a kind of irrationality. Since CONDITIONALIZATION is a rule about how credences should evolve over time, defending it in this way requires a diachronic Dutch Book—the bets that constitute the book must be offered at different times (since only someone who violates PROBABILISM is vulnerable to a Dutch Book if the bets are offered at the same time). Dutch Book arguments, especially diachronic ones, are not as highly regarded as they once were.[5] So I do not regard van Fraassen's defense of CONDITIONALIZATION as compelling. But I am not here trying to decide that question. I am interested in whether Bayesians who do accept CONDITIONALIZATION can endorse IBE.

[5] See Christensen (2004); Hájek (2008a,b).

It is important that the case van Fraassen presents is one in which there is an intuitively obvious prior probability distribution, and given that we can straightforwardly calculate what our posterior probabilities should be when we get different evidence, if we conditionalize. And, moreover, it is a case in which it is very hard to deny that we *should* conditionalize, since we can straightforwardly calculate obvious conditional probabilities. In this way, it is similar to COIN CASE and CARD CASE from my chapter 1. The case van Fraassen constructs involves tosses of a die of unknown bias, and in my view is somewhat unnecessarily complex. We can make similar points using my COIN CASE.

In COIN CASE, I claim that after n consecutive heads flips, your credence $cr_n(H)$ should equal $\frac{1+2^{2n+1}}{2+2^{2n+1}}$. And this is so because that is what your prior credence in H should be conditionally on n consecutive heads flips without a tails flip. Now, van Fraassen assumes that proponents of IBE will endorse something like the following rule.

> BONUS POINTS. If S has $cr_K(H|E) = x$ and H is a sufficiently "lovely" explanation of E, when S learns E she should update her credences so that $cr_{EK}(H) > x$.

If BONUS POINTS is how IBE works, then since the hypothesis that the coin is biased towards heads is obviously a much lovelier explanation of, say, 10 consecutive heads flips than the hypothesis that it is fair, BONUS POINTS would entail that we should have $cr_n(H) > \frac{1+2^{2n+1}}{2+2^{2n+1}}$. I agree that this is wrong.

However, the beginning of any response to van Fraassen's argument is to notice that IBE as I have defined it does not entail anything like BONUS POINTS. So the friend of IBE can accept that when one can precisely calculate a conditional probability for H given E from a reasonable prior probability distribution, then one *should* conditionalize. Under those circumstances, one ought to have a posterior credence equal to one's (obviously justified) prior conditional credence. Under these conditions, the best explanation will just be the likeliest explanation. But these circumstances are special, and the point may not be generalizable. In particular, one feature of COIN CASE is that we know the objective chances of various outcomes. Given that, we are required to have the credences I claim we should even if we do not assume CONDITIONALIZATION. OBJECTIVE CONDITIONALIZATION and the Principal Principle are enough, since the latter tells us that all permissible priors are such that $cr_\emptyset(H|E\langle ch(H|E) = x\rangle) = x$. So the case is insufficient to demonstrate that we should always prefer conditionalization to BONUS POINTS, even.

Moreover, it is clear that in COIN CASE we do infer the loveliest explanation. It may be that what makes it lovely is that it is likely, so that loveliness and likeliness are not really distinct, and IBE is not interestingly different from CONDITIONALIZATION. But neither COIN CASE nor van Fraassen's case provides a counterexample to the claim that we ought to infer the loveliest explanation.

Nevertheless, it is clear that BONUS POINTS conflicts with CONDITIONALIZATION, so if we should *ever* obey bonus, we must reject CONDITIONALIZATION. Some defenders of IBE have gone so far as to claim that we should *never* adjust our credences as BONUS POINTS requires, though we should still take loveliness to be a guide to truth that is potentially independent of likeliness. This is the compatibilist view, which I will also call (as its proponents often do) the heuristic view.

3.2.1 COMPATIBILISM: THE HEURISTIC VIEW

While Bayesians take PROBABILISM and CONDITIONALIZATION to provide normative constraints on ideal reasoning, they do not necessarily assume that actual human reasoners live up to these ideals.

> But some cognitive psychologists have also objected to Bayesianism on the interestingly different grounds that it fails to capture our *bad* inferential practices, and so fails to capture the way we actually reason (e.g., D. Kahneman and Tversky (1982)). What they have done is to construct inferential situations where a Bayesian calculation would yield what is clearly the correct answer, yet most people give a different and incorrect answer.... This pulls apart the normative and descriptive issues. It is here assumed that Bayesianism dictates the inferences one ought to make, but argued on empirical grounds that it does not describe the inferences that people actually make. It is thus inadequate as a solution to the descriptive problem of induction, and the door is opened to the claim that Inference to the Best Explanation does better, descriptively if not normatively. (Lipton, 2004, 105)

Thus, where van Fraassen (1989) constructs a case in which CONDITIONALIZATION is obviously the right response to the evidence, it is open to the compatibilist to say that although in this case BONUS POINTS would lead us astray, BONUS POINTS is an incorrect description of the role that explanatory considerations play in determining our credences. In the view of compatibilists such as Lipton (2004), Okasha (2000) and McGrew (2003), IBE is a useful *heuristic* for approximating ideal reasoning when we do not have precise enough information to employ Bayes' theorem explicitly. So compatibilists reject BONUS POINTS and endorse CONDITIONALIZATION. But they maintain that there are nevertheless cases in which we do reasonably infer the best explanation, where the best is not just equivalent to the likeliest.

> Bayesian conditionalization can indeed be an engine of inference, but it is run in part on explanationist tracks. That is, explanatory considerations may play an important role in the actual mechanism by which inquirers "realize" Bayesian reasoning. As we will see, explanatory considerations may help inquirers to determine prior probabilities, to move from prior to posterior probabilities, and to determine which data are relevant to the hypothesis under investigation. (Lipton, 2004, 107)

Unfortunately, defenders of the heuristic view tend not to be careful about the distinction between different interpretations of the probability function; it is often not clear whether they are talking about subjective credence, objective chance, evidential probability, or something else. They typically write as though when we do not have information about what the *chances* are, we cannot conditionalize. But of course, one might be ignorant about the chances while still having well-defined subjective credences. I will argue that more careful attention to these distinctions undercuts the claim that there is an important role for IBE to play in our inferential practices, given CONDITIONALIZATION. So if there's any interest to IBE, it will be because at least sometimes inferring the best explanation should win out over CONDITIONALIZATION.

If we consistently treat the heuristic view as making claims about subjective credence, then it says that sometimes we mere humans find that we do not *have* well-defined credences $cr_K(H)$, $cr_K(E)$, or $cr_K(E|H)$ (perhaps we have not considered H to be a possibility, and so have implicitly assumed $cr_K(H) = 0$, which makes $cr_K(E|H)$ undefined, or perhaps we never anticipated the possibility of getting evidence E and so have $cr_K(E) = 0$, similarly causing $cr_K(H|E)$ to be undefined). In the slogan of Okasha (2000), IBE can illuminate where Bayesianism is silent.

> In those cases where agents respond to new evidence by inventing new hypotheses, the Bayesian model is silent. But IBE provides a useful, if schematic account of what is going on: the agents are trying to explain the new evidence. They think that the best, or perhaps the only, explanation of the evidence lies outside the space of possibilities they have previously considered, so rather than conditionalizing, they invent a new hypothesis. (Okasha, 2000, 707)

After all, CONDITIONALIZATION tells us what to do only when we have well-defined conditional prior credence. If we hit upon a hypothesis in the context of discovery that we previously gave no credence, because we had never considered it, we may use explanatory considerations as a guide to the correct credence we should now give it, and we will choose explanatory hypotheses that predict the evidence, i.e., we will choose them so that $cr(E|H)$ is high, even though possibly this conditional credence was previously undefined.

Or, alternatively, when we do infer H because it is a lovely explanation of E, we will as a matter of fact have $cr_K(H|E) > cr_K(H'|E)$ for all alternatives H', so that conditionalization will agree with our inference to the best explanation. For this latter possibility to be interesting we must assume that we do not have easy access to our own credences, or we would not need IBE—we could just figure out what our credences are and conditionalize. McGrew (2003, 556) makes this explicit: "[compatibilists] must show that in some sense the [explanatory] virtues are more accessible, logically or phenomenologically, than priors and likelihoods." But that assumption is quite plausible—if I have a well-defined credence in the proposition that France will win the next World Cup, I do not know what it is and couldn't easily find out. I know it is lowish, and non-zero, but beyond that I cannot say. I am not the gambling type. This is crucial, because one of the main objections to the heuristic view is that it robs IBE of all interest. If CONDITIONALIZATION is the fundamental norm of inference, then IBE is unnecessary or useless at best, and misleading at worst.

The defense of the heuristic view, therefore, has three parts.

> An articulation and defense of Inference to the Best Explanation might proceed in three stages: identification, matching, and guiding. First we identify both the inferential and explanatory virtues. We specify what increases the probability of a hypothesis and what makes it a better potential explanation; that is, what makes a hypothesis likelier and what makes it lovelier. Second, we show that these virtues match: that the lovelier explanation is the likelier explanation, and vice versa. Third, we show that loveliness is the inquirer's guide to likeliness, that we judge the probability of a hypothesis on the basis of how good an explanation it would provide. (Lipton, 2004, 121)

The heuristic view has it that Bayesianism gives us a complete story about the "inferential virtues," and then seeks to show how these correlate with explanatory virtues. That is, on the heuristic view, the most virtuous explanation is the likeliest explanation, and the explanatory virtues will explicable in terms of probabilistic relations. So, for example, McGrew (2003) argues that the ratio $\frac{p(E|H)}{p(E)}$ captures the "explanatory power" of H as an explanation of some evidence E, because this ratio will be high just in case E is surprising but H renders it unsurprising (if we think of surprisingness as low prior subjective probability). But it is also correlated with high posterior probability, since unless the prior probability of H is extremely low, a high explanatory power ratio will make $p(H|E)$ high as well. So McGrew identifies a feature of virtuous explanations—they render surprising evidence unsurprising—and matches it to a probabilistic relation correlated with high posterior probability.

The last step is to show that this explanatory power ratio is somehow more accessible than the prior probabilities $p(E|H)$ and $p(E)$, so that this aspect of explanatory virtue can serve as a heuristic that helps us approximate CONDITIONALIZATION, which the heuristic view grants is the fundamental norm of inference. McGrew defends this claim with an example.

> An example makes this plain. At a carnival poker booth I espy a genial looking fellow willing to play all comers at small stakes. The first hand he deals gives him four aces and a king, the second a royal flush, and indeed he never seems to come up with less than a full house any time the cards are in his hands. Half an hour older and forty dollars wiser, I strongly suspect that I have encountered a card sharp. I have made no attempt to compute the odds against his obtaining those particular hands on chance; I may not even know how to do the relevant calculation. Nor do I have any clear sense of the probability of his getting just *those* hands given that he is a sharp. For neither $p(E|H)$ nor $p(E)$ am I in a position to estimate a value within, say, three orders of magnitude; the best I can say in non-comparative terms is that each of them is rather low. But I know past reasonable doubt that the explanatory value of my hypothesis is very great. (McGrew, 2003, 560)

Unfortunately, this example does not make it plain to me. McGrew appeals, in his example, to ignorance about objective chances, not about subjective credence. Having made no attempt to calculate the odds shows only that McGrew does not know what his credences $cr_K(H|E)$ and $cr_K(E)$ *should* be, not that they are not well-defined or even somewhat accessible. Nevertheless, I think it is plausible that he does have at least a higher degree of confidence that his subjective credences are such that $cr_K(E|H) > cr_K(E)$ than he does that these prior probabilities take any particular value.

However, to assess what we should do when an explanation is good in this way, we need to also consider what happens in cases where the most explanatorily powerful explanation is *not* also likeliest. If we suppose that what McGrew's sense of the explanatory power of H is matched to is his *subjective* explanatory power ratio, then the following sort of case should also be possible.

> S watches McGrew play poker against the card sharp. She considers the card sharp hypothesis. She also considers a different hypothesis, L, which says that the carnival poker player is *lucky*. That is, she believes that some individuals are blessed with a special invisible property, *luck*, which makes it the case that beneficial outcomes that would have low probability for individuals without luck occur for them at unusually high rates. She assigns low probability to the proposition that any particular person is lucky, but is quite sure that some people are. And, in fact, S thinks that E is somewhat more likely on hypothesis L, because she is skeptical that even the best card sharps could pull off the feats this carnival worker demonstrates. So she has $cr_{K'}(E|H) < cr_{K'}(E|L)$.[6] So for S, the hypothesis L will have greater subjective explanatory power. And, unless she regards H as antecedently more likely that L, L will get higher credence than H when she conditionalizes. But she might also think that lucky people are more common than card sharps.[7]

This case shows that, possibly, if the better explanation is the one with a higher explanatory power ratio, the better explanation will be L. But L is clearly not as a good an explanation. On the other hand, if McGrew were to say that H is better even for S, then if S believes H because it is a better explanation (as it seems she ought to), she'll violate CONDITIONALIZATION by failing to infer the subjectively likeliest explanation—and this example will be a case of conflict in which IBE should win. Now, I think that this is a case in which S should not conditionalize and should instead give higher credence to the obviously better explanation. The problem for the heuristic view is that its defenders cannot say that. If explanatory goodness is a matter of our own subjective probability assessments, which explanation is the best will be relative to possibly arbitrary features of our prior credence distributions. If so, then any defense of why we should (sometimes) infer the best explanation will be parasitic on a defense of CONDITIONALIZATION.

Recall that the heuristic view is committed to the claim that IBE is not *just* conditionalization. If it is just that the explanation with the highest conditional probability on

[6] K', rather than K, because S's background knowledge will be different than McGrew's.

[7] McGrew (2003) also offers an account of the explanatory virtue of "consilience," i.e., the ability to explain multiple, unrelated pieces of evidence. While I do not think I need to get into a discussion of consilience to make the general point in this section, note that an appeal to McGrew's probabilistic explication of consilience will not necessarily give us reason to favor H over L, either. For in fact the luckiness of the carnival poker player has a greater potential to explain unrelated facts about his life that have nothing to do with cards. This may make it less likely to hold up on examination of additional evidence, but that is precisely why we like explanations with a high degree of consilience, and cannot be a reason to prefer H without the weight of that additional evidence.

the evidence always strikes us as loveliest, or is always the most virtuous, then IBE loses its interest. Thus, it is possible that IBE can lead us astray. In the example above, if S goes astray,[8] it must be because she actually has a sufficiently high prior credence in H to overcome the higher explanatory power of L and give H a higher conditional credence on E, so that if she infers L because it is better in terms of explanatory power, she'll go wrong by the lights of CONDITIONALIZATION. But this is not the way that we usually think that explanatory reasoning leads us astray. Consider the following example of bad explanatory reasoning.

> My second example of a lovely non-causal explanation concerns reward and punishment, and is based on the influential work in cognitive psychology by Daniel Kahneman and Amos Tversky (D. Kahneman and Tversky, 1982, 66–68).... Flight instructors in the Israeli air force had a policy of strongly praising trainee pilots after an unusually good performance and strongly criticizing them after an unusually weak performance. What they found is that trainees tended to improve after a poor performance and criticism; but they actually tended to do worse after good performance and praise. What explains this pattern? Perhaps it is that criticism is much more effective than praise. (Lipton, 2004, 32)

In later chapters, Lipton refers back to this as an example of a case in which our heuristic use of IBE leads us astray (but in a way that is correctable by IBE as well, when we realize that the null hypothesis and regression to the mean is actually a lovelier explanation). But how does it lead us astray? If E is the evidence described in the case and H is the hypothesis that criticism is more effective than praise, this hypothesis will have explanatory power in McGrew's sense if we have subjective credences such that E is surprising and H renders E relatively unsurprising. In fact, this is the reaction that many people have, suggesting that they do give E low prior credence. The mistake is not that they infer H because it is a powerful explanation of E even though they give H very low prior credence and so violate CONDITIONALIZATION. Instead, the mistake is that they fail to give E an appropriate prior credence—because of the phenomenon of regression to the mean, we should have a relatively high prior credence in E, i.e., it should not be surprising. So the mistake is that we have priors that we should not have, when we infer H because it is a powerful explanation of E in this sort of case. That is to say, if H is the likeliest explanation of our evidence, given

[8]I mean go astray by the lights of CONDITIONALIZATION. Again, I think that in fact S should not conditionalize.

our subjective priors, it is because our priors are bad. But this suggests that we should not conditionalize, under the circumstances.

I think this is the *right* way to describe the mistake in the case. On the assumption that one should *not* infer H in this case and should instead revise one's judgment about the extent to which E confirms H, this is a case in which both CONDITIONALIZATION and inference to the loveliest explanation are inappropriate (if one finds H to be a lovely explanation). But the heuristic view has some room to defend the claim that this is a case in which we should infer the best explanation. The defender of the heuristic view may say that it is true that one should not have given E low prior probability, but this is because explanatory considerations should have influenced one's priors through the two mechanisms that the heuristic view allows. For recall that one of the roles the heuristic view envisions for IBE is establishing initial probability assessments in cases where the prior probability of some proposition is not well-defined. Thus, if explanatory considerations play a role in estimating probabilities for previously unconsidered propositions, the effects of those explanatory considerations will shape our later credence distributions.

> Another obvious place to look for a way explanatory considerations might in practice play an important role in a Bayesian calculation is in the determination of prior probabilities. I begin with a general observation about the role of priors in Inference to the Best Explanation. Choices between competing potential explanations of some phenomenon are often driven by judgments of which of the explanation has the highest prior. This is one important source of suspicion about Inference to the Best Explanation: the choices here seem actually to be based on judgment of which is the likeliest explanation, judgments which in many cases depend on which potential explanation is judged to have the highest prior, not on which would be the loveliest explanation (Salmon, 2001, 83–84). My reply is to agree about the crucial role that priors play in this way, but to deny that this is in tension with Inference to the Best Explanation. Consider what the Bayesian himself says about the priors. He of course does not take their crucial role to undermine the importance of the Bayesian formula, in large part because today's priors are usually yesterday's posteriors. That is, the Bayesian claims that today's priors are generally themselves the result of prior conditionalizing. Similarly, the defender of Inference to the Best Explanation should not deny that inference is mightily influenced by the priors assigned to competing explanations, but she will claim that those priors were themselves generated in part with the help of explanatory considerations. (Lipton, 2004, 115)

Similarly, the defender of the heuristic view may say, about my character who believes in *luck*, that she should not give high prior credence to L because it lacks some explanatory virtues that should have caused us to give it low credence whenever we first considered it, or it will have already been disconfirmed by the evidence. On this view, the primary method by which explanatory considerations influence our judgments of posterior probability is through influencing our priors.

However, if the heuristic view is committed to CONDITIONALIZATION as the fundamental norm of credence revision, then the heuristic view's claims about the priors we should have had does not have the right sort of force to require a change in our posterior credences. The posterior credence $cr_{EK}(H)$ was arrived at via conditionalization, and now learning that $cr_K(E)$ should have been higher has no clear relevance to H. We might thereby come to see that H is not, in fact, a particularly lovely explanation of E. But so what? A defender of IBE should want to say that this should cause us to revise *downward* our credence in H, but if explanatory loveliness is trumped by conditionalization when we can conditionalize, then the fact that considerations of loveliness would have caused us to give E a higher initial credence the first time we considered it (had we been aware, at the time, of regression to the mean) should not matter given the fact that the posterior credence we have now was arrived at by CONDITIONALIZATION. That is, even if explanatory considerations do help us to fix initial credences when they are ill-defined, unless IBE should cause us to change our posterior credences, it will be toothless. Lipton's air force case is one in which IBE can both lead us astray and get us back on track, but the heuristic view seems to require that it can *only* lead us astray; once we've started on the wrong path, we have to conditionalize. Otherwise, we get cases of conflict between likeliness and loveliness that should be revised in favor of loveliness. The heuristic view is therefore not satisfying, if we think IBE does have some important role to play in our inferential practice.

3.2.2 OBJECTIVE BAYESIAN IBE

The heuristic view falters because it robs IBE of any normative force apart from that of CONDITIONALIZATION. But CONDITIONALIZATION is already suspect, because sometimes one might have impermissible priors, and in such a situation one should not conditionalize. Weisberg (2009) defends an objective Bayesian take on IBE as a way of remedying this defect. That is, rather than endorsing CONDITIONALIZATION, Weisberg argues that we should endorse OC, repeated here for ease of reference.

> OBJECTIVE CONDITIONALIZATION (OC). For all propositions P, and subjects S, S ought at t to have $cr_t(P) = o(P|E)$, where o is a permissible prior probability distribution, and E is S's total evidence at t.

Thus, when the likeliest explanation in terms of our *subjective* credences is objectively bad (i.e, no permissible priors give it low credence conditional on our total evidence), we should not conditionalize, but should instead adjust our credences so that our posterior credence agrees with those of a permissible prior, conditioned on our evidence. In the view of Weisberg (2009), there is a unique permissible prior probability distribution, which I will again call the "One True Priors." Since Weisberg makes this assumption (though possibly just for simplicity), I shall do so as well in the rest of this section.

This strategy for fitting IBE into an objective Bayesian framework shares some key features with the heuristic approach. It still endorses a fundamental norm of credence revision that makes no reference to explanatory virtues or loveliness, and then seeks a role for IBE by matching virtuous explanations to explanations with high conditional probability by the lights of the One True Priors.

This objective Bayesian take on IBE gives the right result in the Israeli air force case. For the mistake there is again that people tend to have an inappropriately low prior credence in E, given their background knowledge K. But the fact that it is inappropriate shows that $o(E|K)$ is higher, and so $o(H|EK)$ is lower. OC thus requires that we adjust our posterior credence in H downwards, even without any additional evidence. OC has the right sort of normative force to require an adjustment to our posterior credence even if it was arrived at via CONDITIONALIZATION. It also gives the right result in the card sharp/luck pair of cases. There, presumably $o(L|E) < o(H|E)$, so S should not give higher posterior credence to the lucky hypothesis L. This is the main point in favor of Weisberg's objective Bayesian IBE, and I agree that it is a virtue.

An interesting feature of this account is that loveliness and explanatory virtue play no role in the explanation of the correct credence adjustments in these cases, and seem altogether to disappear from the story. Weisberg's IBE is inference to the likeliest explanation by the lights of the One True Priors. This may not be the explanation that *strikes us* as loveliest.

One elegant way of fitting loveliness back into the story would be to *accept* the heuristic view's claims that loveliness matches with high subjective conditional probability, while matching explanatory virtue to probabilistic relations between propositions that determine a high objective conditional probability. This would yield a nice explanation of the Israeli air force case. On this view, what we would say about that case is that inferring the *loveliest* explanation leads us astray because CONDITIONALIZATION does, while inferring the *most virtuous* explanation gets us back on track to adhere to OC.

It is not perfectly clear from the text whether Weisberg would endorse this account of explanatory loveliness. But his sales pitch is that friends of IBE should endorse OC, because it, unlike the heuristic view, allows IBE to play a substantive normative role in our inferential practice. On the story about loveliness I just told, it is easy to see why this might be unsatisfying. If friends of IBE want to defend the claim that inference to the loveliest explantion is a good inferential practice, this story fails to deliver the desired result. For on this objective Bayesian story, IBE just is OBJECTIVE CONDITIONALIZATION. Although it gives more intuitively correct results in some cases, it seems that this story nevertheless fails to preserve any particularly interesting role for inference to the loveliest explanation. OC gives the complete story about how our credences should evolve, and explanatory loveliness plays no constructive role (it is at best superfluous, at worst misleading to infer the loveliest explanation).

Moreover, on this story it is hard to see what's special about *explanatory virtue*, either. There are indications that Weisberg wouldn't endorse this equation of explanatory virtue to high conditional probability on the evidence; he's certainly sensitive to the charge that explanatory virtue needs to be distinct from high conditional probability to be interesting, as he levels this charge agaist the heuristic view.

> Of course, one can always rig one's account of explanatory goodness so that compatibility is guaranteed. If "best" just means "has the highest prior conditional credence," then compatibility with [CONDITIONALIZATION] is a no-brainer. The compatibilist who gives the above account of explanatory goodness certainly is not guilty of anything so egregious. She identifies common-cited explanatory virtues with probabilistic ones—fit to background belief with $p(H)$, and fit to the data with $p(E|H)$—and then relies on Bayes's theory to bind explanatory goodness to prior conditional credence. But the question remains whether the notion of goodness she offers is thick enough to preserve the spirit of IBE.

> I think the answer is clearly negative. On her account, fit with background belief and fit to the data are the only contributors to explanatory goodness, leaving out simplicity, elegance, unification, and the rest of the standard explanatory virtues. In fact, there is nothing intrinsic to the hypothesis or its relationship to the evidence that is relevant to explanatory goodness, on her picture. Explanatory goodness is, for her, *entirely* a matter of the agent's degrees of belief. (Weisberg, 2009, 132)

A similar complaint could be leveled at the objective Bayesian account of IBE if it equates explanatory virtue with probabilistic relations that determine high conditional probability (by the lights of the One True Priors). If we make this equation, then explanatory virtue is just a matter of high evidential probability. That is, possibly an objective Bayesian IBE says that the permissible priors are what they are for reasons that have nothing to do with explanatory considerations, and what we call "explanatory virtues" are just features of this probability function—possibly H is a better explanation of E than H' just in case $o(H|E) > o(H'|E)$, but any reasonable prior function has this feature independently of whether we find H to be a lovely explanation, or whether H is simpler, more elegant, etc. This does not preserve any interesting, independent normative role for explanatory considerations. So in the spirit of the quotation above, the objective Bayesian defender of IBE should endorse the following.

> Difference. Possibly H is a more virtuous or lovelier explanation (for S) of E than H' given background assumptions K, but there are (otherwise reasonable) probability functions p and p' such that $p(H|EK) > p(H'|EK)$ and $p'(H|EK) \leq p'(H'|EK)$.[9]

difference entails that being a more virtuous or lovelier explanation does not necessarily determine that the virtuous/lovely explanation will also have higher conditional probability on the evidence for *just any* reasonable probability function.

The objective Bayesian proposal then amounts to the claim that despite this, explanatory considerations have an important role to play in determining which priors are *epistemically permissible*.

> Indispensibility. Necessarily, for all subjects S, hypotheses H, H', evidence E, and background knowledge K, if H is a lovelier/more virtuous explanation of E (for S) given

[9] By "otherwise reasonable" I mean at least that neither p nor p' are impermissible by the lights of the Principal Principle, or other relatively uncontroversial constraints on permissible priors, but which do not themselves depend on explanatory considerations to rule out certain priors as impermissible.

K, then S epistemically ought to have $cr_\emptyset(H|EK) > cr_\emptyset(H'|EK)$.[10]

INDISPENSIBILITY and DIFFERENCE entail that facts about what the One True Priors are determined by explanatory considerations.[11] So we should infer the most virtuous/loveliest explanation not simply because doing so happens to coincide with OC, but because which prior probability distributions are permissible is a function of explanatory considerations. This suffices to preserve an important normative role for explanatory considerations to play in our reasoning even if we should always obey OBJECTIVE CONDITIONALIZATION.

Assuming DIFFERENCE, there's also room for a more "heuristic" version of objective Bayesian IBE. If we assume there there is a unique permissible prior probability distribution o, then it might well be that *usually* when H is a better explanation of E than H' (given K), $o(H|EK) > o(H'|EK)$, but not always. It is fairly obvious that explanatory loveliness, especially, if not explanatory virtue, is much more accessible than objective evidential probability. So it might be reasonable to use explanatory considerations as a heuristic guide in trying to live up to OC, without taking explanatory considerations to have the strong role that INDISPENSIBILITY grants them. A potentially oversimplifying way of representing this might be as follows.

HEURISTIC. When H is a lovelier/more virtuous explanation of E than H', for a subject S with background knowledge K, S ought to give high credence to the proposition that $o(H|EK) > o(H'|EK)$.

The oversimplification is taking explanatory considerations to confirm the proposition expressed by the inequality, which, surely, is not a proposition one generally entertains, let alone endorses, even when engaging in explanatory reasoning. But, since presumably when we do have evidence about what the One True Priors are, we will (if we are rational) make sure our priors conform to what we believe about the contours of this function, the proposal

[10] One may quite reasonably worry here that the quantification over subjects here will lead to an overly permissive take on the One True Priors, at least if it is explanatory loveliness that is supposed to be indispensible. I will address this worry in §3.3.1 below.

[11] Weisberg does not quite endorse DIFFERENCE and INDISPENSIBILITY, but I think they are in the spirit of his proposal. "But despite its limitations and underdeveloped state, the proposal serves its illustrative purpose: it demonstrates how explanationism might be fit into an objective Bayesian framework, helping to constrain a priori probabilities. In a situation where B, H, and E are such that the Principle of Indifference applies but does not yield a unique value for $p(H|EB)$, we can use [explanatory considerations in the way] just outlined to calculate (or maybe just further constrain) $p(H|EB)$" Weisberg (2009, 19–20).

is that insofar as explanatory considerations are a defeasible guide to the credences we ought to have, we can represent this as gaining information about what our priors ought to be.[12]

This heuristic version of objective Bayesian IBE is not open to the main line of objection against the subjectivistic take on the heuristic view. For, as noted above, it does oblige revision of posterior credence in cases like the air force case, even if it is not IBE itself that obliges us to make such revisions. It may still be open to complaints that it does not preserve enough of a role of IBE, whose normative force is entirely subordinate to that of OC. A further reason to doubt the heuristic view is that there may be no reason to think that there is a One True Prior probability distribution that is specifiable independently of explanatory considerations. Objective Bayesians do not typically posit such a unique function, and even the principles that objective Bayesians have proposed as rational constraints on prior probabilities do not always yield unique answers, even in the cases to which they are meant to apply. For example, the Principle of Indifference, which enjoins us to treat as equiprobable members of a partition of the probability space that we have no reason to treat differently, cannot apply independently of a particular way of dividing the space of possibilities. Thus, Weisberg (2009, 137) hopes that his proposal can "maybe even solve some of the problems that have hampered the objective Bayesian program," particularly by giving some guidance about what our priors should be when there is more than one possibly correct way to divide up the space of possibilities. For this reason, I take it that Weisberg means to be proposing a strong take on IBE.

We can also distinguish between what appears to be Weisberg's preferred view, which takes inference to the most virtuous explanation to provide the relevant norm, and the more subjective inference to the loveliest explanation. Of course, loveliness may generally be a guide to explanatory virtue, so these may not be so distinct. However, since it is not at all clear just what the explanatory virtues are or how to characterize them, or if a non-subjective characterization of explanatory goodness is even possible, in the rest of this chapter I will develop and defend a subjective take on IBE that relies on the subjective notion of *intellectual satisfaction*. That is, I will defend the claim that when H is the most intellectually satisfying

[12] In a similar fashion, we can quite generally represent evidence as information about probability functions to which we give varying degrees of deference; the One True Priors will be an "expert" probability function to which we give complete deference. Cf. Joyce (2011).

explanation of some evidence E, given background knowledge K, then one ought to have $cr_K(H|E) > cr_K(H'|E)$ for all alternative explanatory hypotheses H'.

3.3 INTELLECTUAL SATISFACTION AND INFERENCE TO THE BEST EXPLANATION

White (2005b) defends the limited thesis that explanatory considerations guide our enumerative inductive reasoning.[13] In the process, he goes some way towards offering a characterization of what the explanatory value of a hypothesis consists in, a characterization that I largely agree with.

> I would suggest that the relevant notion of explanatory value has to do with the degree of *satisfaction* that an explanation should deliver, if assumed to be correct. In asking a why-question we are seeking to satisfy a pecular kind of curiosity; we are seeking *understanding* and trying to *make sense* of things. The kind of satisfaction that a good explanation can deliver does not simply consist in our taking ourselves to have discovered the truth. For even if known to be correct, an explanation can remain deeply unsatisfying. . . .
>
> The way then to assess the value of an explanation is to ask: Suppose we were to learn for certain that this explanation is correct. How satisfying should it be in our quest for understanding and making sense of things? On this conception of explanatory value, to answer the question "How good an explanation does this hypothesis provide?" is not simply to judge its *plausibility* as an explanation. For we are factoring out the question of its truth by considering how satisfying it should be, *if* we were to know that it is true. (White, 2005b, 2)

This account of explanatory value has several features worth highlighting.

- Accessibility. We can judge whether a hypothesis has explanatory value by reflection.
- Hypotheticality. We can judge whether a hypothesis has explanatory value independently of whether we have the evidence it explains and independently of our subjective credence in either the explanans or the explanandum.
- Subjectivity. Whether or not a hypothesis has explanatory value depends on whether it satisfies our curiosity.

[13] "I wish to leave open the question of what role, if any, non-explanatory considerations should play in epistemic evaluation. My aim is simply to show that explanatory considerations have an important role to play in the evaluation of inductive hypotheses" (White, 2005b, 1).

This last item especially requires some explanation. White (2005b) endorses accessibility but not necessarily in a way that obviously supports the subjectivity of explanatory value.

> Perhaps in many cases we can judge explanatory value without the aid of explicit guidelines. But it can be illuminating to consider what are some of the factors that make for a good explanation. (White, 2005b, 2)

This suggests a picture on which the *felt satisfaction* of our curiosity which is a marker of explanatory value is merely the subjective correlate of more objective features of the explanatory hypothesis. The particular feature White (2005b, 4) appeals to is:

> STABILITY. An explanation of a fact F is *stable* to the extent that according to this explanation, F couldn't easily have failed to obtain.

Stability is plausibly a feature that can be explained in terms of objective counterfactual relations between the evidence to be explained and the explanatory hypothesis, and it therefore may seem that explanatory virtue is not just a function of the satisfaction of our curiosity.

However, stability is not necessarily an explanatory virtue. Often, perhaps, stable explanations are more satisfying, but stability alone does not explain why.

> STABILITY CRITERION. Stability is a virtue of an explanation to the extent that its explanandum calls for an explanation. (White, 2005b, 4)

White (2005b) does not offer a thorough account of when something calls for an explanation, but does offer the following necessary condition.

> SALIENCE CONDITION. If F is a member of a homogenous partition of G, and G was bound to obtain, then F does not call for an explanation.
>
> A *homogenous* partition of G is a class of states of affairs (pair-wise inconsistent, and whose disjunction is equivalent to G) such that no member of the class stands out as *any more* in need of explanation than any other would, had it obtained. (White, 2005b, 3)

This condition itself appeals to the notion of explanatory urgency in the characterization of homogenous partitions. To see this, consider that if we flip a coin 100 times, the 2^{100} possible outcomes do *not* form a homogenous partition. For a sequence of 100 heads flips would stand out as much more in need of explanation than any sequence that contains roughly 50 heads and 50 tails (and as the ratio departs from 1:1, the explanatory urgency will presumably increase). So we cannot just appeal to partitions of equiprobable events one of which is bound to obtain, in the explanation of this necessary condition on explanatory urgency. Otherwise, a sequence of 100 heads would not need explanation.[14] Stable explanations are satisfactory only if they explain phenomena that activate our curiosity, so stability is not a sufficient condition for explanatory value. Nor is it a necessary condition; the hypothesis that the coin is fair and the outcomes random is a satisfactory explanation of any sufficiently random sequence of heads and tails results with a ratio close to 1:1, despite the fact that this explanation does not make that particular outcome at all likely, and many other outcomes could have easily occurred instead, given that hypothesis. So for all White (2005b) says about it, at any rate, explanatory value is a subjective notion in the sense that whether a hypothesis has explanatory value just depends on whether it satisfies our curiosity—whether we find it intellectually satisfying, or not. I do not necessarily expect that White would endorse the subjective version of IBE that I will defend; my point here is that it is consistent with White's examination of some of the features that make for a good explanation.

Thus, while I do not think investigation into features common to satisfying explanations is pointless, I am not optimistic that we will be able to identify objective features of explanatory hypotheses that are necessary and sufficient for us to find an explanation satisfying. Moreover, even if we should find a list of features that we all agree make for satisfying explanations, this is no guarantee that we would then have any independent explanation of why those features are correlated with the *truth* of a hypothesis. Nor do I want my defense of IBE to be put off until we have this list of features. So I will defend an account IBE as inference to the most intellectually satisfying explanation, where intellectual satisfaction is to be understood subjectively. That is, H is an intellectually satisfying explanation of E when our curiosity about E is satisfied by H.[15] Thus, when H is the most intellectually satisfying explanation of E against background knowledge K, one infers the best explanation when one has $cr_K(H|E) > cr_K(H'|E)$ for all alternative explanatory hypotheses H'. Note

[14] Possibly, if we knew that the coin was fair, we would have to accept that this is just one of 2^{100} equiprobable outcomes, and there is no further explanation. It seems to me that this would not remove the felt urgency of explanation, though. With any justification short of certainty for the hypothesis that the coin is fair, we ought to suspect that our belief in its fairness is mistaken and look around for other explanations.

[15] Lest it seem that this gives up on the hypotheticality of IBE, let me point out that we can by hypothetically curious. I can certainly ask, of some E that I do not expect to get, what would be a satisfying explanation of it if I did. While perhaps this is often pointless, we can see this operating in reverse by asking, of some hypothesis that we do not believe, what evidence it might take to convince us of its truth.

that on this account of IBE, the name "inference to the best explanation" is something of a misnomer. For one need not infer H at all, either in the sense of believing H, or even in the sense of giving it higher credence. This is a feature of the hypothetical nature of IBE; we can assess the explanatory value of a hypothesis whether or not the evidence to be explained is actually known by us or merely entertained as possible evidence.

However, this is not yet a complete account of the role of IBE; the accessibility and hypothetical nature of IBE forces us to consider cases in which we reflectively judge that H *would* be an intellectually satisfying explanation of E if we add or subtract additional assumptions from our background knowledge K. For example, consider the following case of a situation that urgently requires explanation and some possible explanatory hypotheses.

> To illustrate, suppose I were to levitate off my chair and float across the room and out the window. I should very much like to learn why this was happening. Now, contemporary physics suggests that this is quite possible, although the chance of its happening is extremely low. Randomly moving air molecules may all happen to align and move upwards with enough momentum to lift me. It may be that the correct and most thorough explanation of why this happened appeals to nondeterministic laws which assign it an extremely low probability. As I drift across Manhattan, were I somehow to learn that such an explanation is correct, including every detail of the physical laws, I would hardly respond, "Oh, okay, so that is why I am floating—for a while there I was puzzled." Indeed I might be even more puzzled by this explanation than if I learnt that it was because a witch put a spell on me. The latter explanation would be hard to swallow, since I do not believe in witches. But at least if there *were* witches we might expect this sort of thing to happen. (White, 2005b, 6)

Let P be the nondeterministic story that gives this evidence E a low probability, and W the witch hypothesis. Let us suppose that White knows that there is no such thing as witches.[16] Then $cr_K(P|E) > cr_K(W|E)$, because $cr_K(W) = 0$, and $cr_K(P|E)$ is very very close to zero but non-zero. Nevertheless, White suggests here that he would find the witch hypothesis more satisfying. This calls for additional care in what we call the "best" or "most satisfying" explanation. Suppose that P and W are the only live options in this case;[17] if White gives P

[16] I am unsure whether it is more charitable to read White as knowing that there are no witches or not, in interpreting this passage. If Williamson (2000) is right that one's evidence is just what one knows, then if White knows that there are no witches, he should have zero credence in the witch hypothesis. Yet he suggests here that he would regard floating as evidence that confirms it; if so, then he should not give it zero prior credence, since even if P is true (and he gives non-zero credence to P), he expects that there is a slight chance that he could get evidence E. But one should not now give zero credence to hypotheses that one thinks one might get evidence for in the future. So I do not claim that the following reading of this passage would be endorsed by White, but I will nevertheless use his example to illustrate my point.

[17] It might be that it is not compatible with a possibility having probability zero that it count as a live hypothesis for purposes of IBE; if so, I can say instead that what we should say about this case is that if W

higher credence conditional on E because $cr_K(W) = 0$, shall we say that he infers the best explanation? Or does he fail to do so, since W *would* be more satisfying if it were true?

In fact it seems to me that I would not find the witch hypothesis satisfying, were something like this to happen to me. Satisfaction, though in principle something we can assess hypothetically, given variations in background assumptions, will of course typically be sensitive to the assumptions that we actually make and do not explicitly cancel. So this is not a case, I think, in which we ought to prefer a less satisfying explanation simply because we give the more satisfying explanation zero credence; rather, it is a case in which W is potentially more satisfying than P given a wide range of background assumptions, but it is not more satisfying given our actual background knowledge. Satisfaction can be hypothetically assessible while being usually assessed non-hypothetically. I will say that an explanation H is actually the best explanation of E for S when H is the most satisfying explanation of E for S given S's actual background knowledge K. So in the case above, inference to the best explanation would endorse belief in the low-chance, relatively unsatisfying explanation P. Given the failure of the luminosity of knowledge[18] one may therefore mistakenly believe that W is a better explanation, precisely because it is a more satisfying explanation of E given some background assumptions K' which differ from what one actually knows (and do not entail that W has probability zero). This is one of the ways that we can be mistaken about what the best explanation of the evidence is. In this kind of mistake, one's hypothetical assessments of explanatory value are reliable, but one fails to have appropriate credences because one conditionalizes one's permissible prior on bad evidence.

I am *not* here saying that we should not reconsider whether there are witches, if we should find ourselves floating above the city. P is not a satisfying explanation, even if W is not either. And as I will claim in § 3.3.2, we should look for stable explanations of surprising phenomena. P is unsatisfying because it is not stable. This might well cause us to reconsider whether W, or to infer that there *must* be an explanation other than P and W that one simply has not thought of. Given that knowledge is not luminous, one might

is a live option, it is more satisfying than P, but W is not a live option given what we know, so in fact P is the most satisfying of the hypotheses under consideration, simply in virtue of running unopposed, as it were.

[18] The principle that when one knows that P one is always in a position to know that one knows that P (Williamson, 2000).

even reasonably seek evidence that one knows that $\neg W$, to make doubly sure that one has not made a mistake—in this sense, we can possibly "confirm" even things that we know are true. But an unsatisfying explanation might still be the *most* satisfying explanation—in the extreme, if E is extremely unlikely, given P, but we can rule out all other possible explanations with certainty, then we should be certain that P nonetheless. Floating out the window and above the city would certainly make one extremely curious about the cause of that phenomenon, and even if the witch hypothesis is inconsistent with what one knows, one certainly ought to try to either come up with a more satisfying explanation than P, or at least to gather additional evidence for P if possible.

We can sidestep this complication by spelling out IBE in terms of one's theoretical prior probability function, $cr_{\emptyset}$, rather than one's current actual credence function given one's background knowledge K, cr_K.

> Inference to the Best Explanation. When H is the most intellectually satisfying explanation of potential evidence E for an ideally rational subject S given additional assumptions K, S should have $cr_{\emptyset}(H|EK) > cr_{\emptyset}(H'|EK)$ for all less satisfying explanations H'.

Given objective conditionalization, this account of IBE entails that one should have $cr_K(H|E) > cr_K(H'|E)$ when H is the most satisfying explanation of E given one's actual background knowledge, even if one in fact mistakenly has $cr_K(H|E) \not> cr_K(H'|E)$ because one finds H' more satisfying by assessing it is explanatory value against mistaken background assumptions K'. Thus, the subjectivity of intellectual satisfaction does not eliminate the potential for distinguishing between the explanation that is actually best and the explanation that strikes us as loveliest. This is so even though the "best" explanation is defined in subjective terms, as the one that would strike us as the loveliest explanation of our actual evidence.

There's another, related way that we can go astray by trying to infer the best explanation and failing to do so. It is important to distinguish the notion of intellectual satisfaction appealed to in my account of IBE from other forms of satisfaction. For example, if someone dear to me were on trial for murder, and the hypothesis that she did it best explains the evidence presented at the trial, I might nevertheless fail to find this explanation satisfying because I have a strong personal desire to believe in her innocence (let us suppose that the rest of my evidence does not strongly conflict with the hypothesis or suggest an alternative, lovelier explanation). That is, I may find the innocence hypothesis more satisfying not because it satisfies my *curiosity*, but because it satisfies some desire of mine other than the desire to understand.

This is not a distinction that White (2005b) makes explicitly, but it is clearly important. The intersubjective variability of desires and interests would render any view that appealed to satisfaction, as opposed to a specifically intellectual satisfaction, quite permissive. For if explanatory considerations should constrain our priors but explanatory considerations are dependent on our desires and interests, then the variability in the latter will entail that I may value H more highly as an explanation of E, even if we otherwise share the same background knowledge. For if I find H satisfying not because it is a better explanation of E than H', but because I prefer that H is true, and if you simultaneously permissibly believe H' because you lack the desires or interests that cause me to prefer H and therefore do not find it more satisfying, then it will be hard to maintain the view that the satisfaction that H provides me as an explanation of E should constrain my credence. For if I know that it is permissible for you to believe H' (or even simply to give it higher credence than H) on the basis of an evaluation of the exact same evidence, then I may reasonably regard myself as permissibly believing H'. At the very least, if I could rid myself of the desire to believe H, I could permissibly believe H'. This is just to say that I can believe H or H' as I please, which is to say that I am not constrained to believe H in anything but the sense that I happen to prefer to believe it. This is no real constraint.

Even after distinguishing between intellectual satisfaction and other forms of satisfaction, though, it is easy to feel a lingering worry that similar considerations can be used to generate a skeptical argument against any view of IBE that depends on subjective responses like having one's curiosity satisfied. For subjectivity invites the possibility of intractable disagreement about what's intellectually satisfying, and this suggests that I have no reason to think that satisfying my curiosity rather than yours is more likely to be a mark of truth. In the next section, I will present such an argument and defend IBE against it.

3.3.1 SKEPTICISM ABOUT INTELLECTUAL SATISFACTION

My subjective take on the notion of intellectual satisfaction of curiosity supports the following premise.

> Disagreement. Possibly S finds H to be an intellectually satisfying explanation of evidence E given additional assumptions K while S' does not.

For perhaps S is simply less curious than S'. I have provided above two ways of explaining failures of inference to the best explanation, but quite possibly neither applies. It may well be that S and S' are quite clear about what the evidence to be explained and all the additional assumptions are, so that we cannot appeal to differences in background knowledge to explain their differing judgments of intellectual satisfaction. And we can also hypothesize that this is really a difference of *intellectual* satisfaction.

The worry is that this entails a permissive epistemology, and that this is problematic. White (2005b), for example, argues against any permissivist account of what's epistemically rational. Below is the particular form that disagreement might be thought to support, given my account of IBE.

> Permissivism. Possibly S should have $cr_{EK}(H) > cr_{EK}(H')$ while S' permissibly does not.

White's main line of argument against permissivism is that, if one permissibly infers H from E (given background assumptions K), but one might permissibly have done otherwise, then the factors that determine whether one takes E to confirm H or not are arbitrary. But we hope that examining the evidence is more likely than not to lead us to true beliefs than arbitrarily selecting doxastic attitudes when confronted with propositions (even consistent attitudes). Given disagreement, it may seem that it is arbitrary whether or not I find H to be a lovely explanation of E given K. But then our confidence that loveliness is a guide to the truth should be undermined, if I could have done just as well to not infer H.

> Supposing [one may permissibly infer H or not], is there any advantage, from the point of view of pursuing the truth, in carefully weighing the evidence to draw a conclusion, rather than just taking a belief-inducing pill? Surely I have no better chance of forming a truth belief either way. If my permissive assumption is correct, carefully weighing the evidence in an impeccably rational manner will not determine what I end up believing; for by hypothesis, the evidence does not determine a unique rational conclusion. So whatever

I do end up believing upon rational deliberation will depend, if not on blind chance, on some arbitrary factor having no bearing on the matter in question. (White, 2005b, 448)

Given my particular view, the worry is that whatever I end up believing upon rational deliberation, given that I obey IBE, will depend on what I find intellectually satisfying. But DISAGREEMENT suggests that whatever causes me to find H more satisfying than H' depends on subjective features that others may lack. When they do not find H satisfying in the same way I do, I will have to suppose either that I am in a privileged epistemic position in which *my* intellectual satisfaction has epistemic weight but others' does not, or that finding H intellectually satisfying does not support higher conditional credence in H in this case.[19] The first option is hubristic; the second appears to undermine IBE.

My response to this line of argument will have two parts. First, I will argue that whether or not one finds H to be an intellectually satisfying explanation of E (given K) is not arbitrary, and at least sometimes opens one to epistemic evaluation, so that DISAGREEMENT does not necessarily support PERMISSIVISM. Second, I will argue that even if some disagreement about intellectual satisfaction is consistent with both parties being perfectly epistemically rational, we nevertheless have reason to be *hopeful* that intellectual satisfaction is a guide to the truth. That is, we can live with some degree of PERMISSIVISM, if the cases in which disagreement undermines IBE do not undermine the hope that we can use IBE to arrive at the truth through more careful investigation.

3.3.2 CURIOSITY AND RATIONALITY

Intellectual satisfaction is the product of our desire to understand; as noted in § 3.3, satisfying explanations put to rest an active curiosity about unexpected or poorly understood phenomenon. The aim of understanding is different from the aim of believing the true. But it is anyway implausible that believing the true (and disbelieving the false) is our sole epistemic aim. Understanding is plausibly epistemically valuable; a person with an encyclopedic

[19] If IBE is a mere heuristic guide, in that intellectual satisfaction is defeasible, *prima facie* evidence about what the independently specifiable One True Priors are, then I will have the third option of thinking that intellectual satisfaction is generally a reliable guide to the truth, but in this case I should not trust it. But I am trying here to defend a stronger take on IBE, according to which explanatory considerations determine our epistemic obligations, rather than simply revealing them.

memory for facts gleaned from Wikipedia but a limited ability to articulate any kind of understanding or explanation of any of them would not be someone to emulate.

This is the reason for the invocation, in my statement of INFERENCE TO THE BEST EXPLANATION, of an ideally rational subject. Imagine a wholly incurious individual who never wonders what the explanation for anything is. A coin lands heads 200 times in a row; she merely shrugs and does not wonder why. You suggest that there is no explanation other than chance; unlikely as it is, it just happened. She finds this perfectly satisfying. I claim that this is irrational. A rational person should be curious; she should have the desire to understand.

As noted above, certain phenomena activate our curiosity, and it is these for which stable explanations are satisfying. It seems to me that there are clear cases in which one *should* be curious about what explains some evidence, and it seems somewhat less obvious but still plausible to me that there are some things about which one should *not* be curious, or at least that it is possible to have an excess of curiosity. The latter would manifest in being dissatisfied with perfectly good explanations. Of course, this is no definition—it would be circular to define IBE as inference to the most satisfying explanation, and then say that explanations are satisfying if they are good. Nevertheless, we can (sometimes) tell the difference between evidence that should activate our curiosity or which demands an explanation and evidence that does not, as in White (2005a).

It is possible that sometimes insufficient or excessive curiosity is a manifestation of one of the two ways that IBE can lead us astray described in § 3.3. For example, one might be excessively curious about why President Obama is not taking a stronger stand in favor of raising the debt ceiling without a long-term deficit reduction deal if one mistakenly takes oneself to know that long-term deficit reduction is not one of his priorities; one might then look around for elaborate explanations in terms of complicated negotiating tactics or believe that he's incompetent. This is to say that whether or not some evidence activates our curiosity is not independent of our background beliefs; if a satisfying explanation is already among them, we may well be incurious, and if our background beliefs are false, what best explains them will not provide understanding. Similarly, one might be insufficiently curious about whether President Obama was born in Hawaii because one is a Democrat and so prefers not to believe things that would give the opposition a political advantage. So what activates our curiosity may be similarly dependent on our desires and inclinations.

In the case of a thoroughly incurious person I described above, it might simply be said that she is disinclined to spend mental effort trying to understand the world, and so fails to ever have her curiosity piqued.

It may well be that not all cases of insufficient/excessive curiosity are explicable in one of these two ways. However, even without these explanations for failures of curiosity, I think the following is quite plausible.

> CURIOSITY. An ideally rational subject is curious about surprising phenomena, where P is surprising to S (at t) just in case S has credences such that $cr_t(P)$ is relatively low.[20]

So to the extent that disagreements about which explanations are intellectually satisfying can be traced to differences in curiosity between subjects, these differences *may* not be arbitrary, and may be subject to epistemic evaluation.

Thus, to derive PERMISSIVISM from IBE, DISAGREEMENT is not enough. We need something stronger.

> DISAGREEMENT*. Possibly, S and S' are both perfectly rational, but S finds H to be a satisfying explanation of evidence E given additional assumptions K while S' does not.

For if one of S or S' is less than ideally rational, then it may be that S or S' should not give higher credence to more intellectually satisfying explanations. Intellectual satisfaction is a mark of truth only for sufficiently curious inquirers.

One possible view, therefore, is that ideal epistemic rationality fixes a unique standard of curiosity. Perhaps any two ideally rational agents will have their curiosity activated by the same evidence, and find the same explanations of that evidence intellectually satisfying. However, I expect that it will be objected that we have no reason to think that this is true. Certainly it seems conceivable that two ideally rational agents might nevertheless differ in

[20]Exactly how low, and whether this varies with context or the practical stakes, I leave open. My claim is just we will sometimes reasonably judge that S epistemically ought to be curious about some P that she did not expect, even if we did expect it, and even if, on reflection, her curiosity is satisfied by noting that she ought to have expected it given her prior evidence (though presumably this latter possibility will not occur for ideally rational inquirers). It may also be true that sometimes people ought to be curious about phenomena that they did expect but should not have, though again, assuming ideally rational agents have permissible priors, this possibility applies only because of a different failure of ideal rationality than a failure to be sufficiently *curious*.

the degree of intellectual satisfaction they get from considering H as an explanation of some E (given K). I agree, and I do not want to defend this strong view. The limited claim that I have defended in this section is that whether a subject is curious and desires to understand why some E is true is not a matter of indifference, epistemically speaking. In the next section, I will go on to argue that even if some degree of permissivism is still plausible given that IBE is inference to the most intellectually satisfying explanation, it is a tolerable degree—we will still have reason to be hopeful that inferring the most intellectually satisfying explanation of our evidence is a good way not just to quiet our own curiosity, but to resolve intersubjective disagreement, as well.

3.3.3 HOPEFULNESS

Hopefulness is a property of sources of justification, defined by Weinberg (2007) as follows.

> Let me stipulate the term hopeful for such sources of evidence: a source of evidence that is not practically infallible is hopeful to the extent that we have the capacity to detect and correct for its errors.

A hopeless source of justification might, by contrast, be relatively but not perfectly reliable, but in such a way that we cannot discover and correct the occasional false belief that it supports. In the context in which Weinberg introduces this terminology, his concern is to attribute hopelessness to the philosophical use of intuitions as evidence in arguments, in contrast with hopeful sources of justification like perception.

> Consider perception again. It is important to note that we have much more to go on in our perceptual practices than reassurance of the on-average high accuracy of our eyes and ears. We also have a well-developed capacity for telling when we are in one of the thankfully rare cases in which our senses deceive us. Each sense modality within itself demonstrates significant intersubjective and intrasubjective agreement: most people see most things in mostly the same way, and their later perceptions are similarly in alignment with their earlier ones. Thus we can use multiple looks in and across perceivers to check each other. If we really want to, we can rely on various sorts of instruments, such as a measuring tape, to double-check what our eyes tell us; though again we would not do so except against a background of very general agreement between our senses and the deliverances of other sources of information. That old saw "measure twice, cut once" recommends to us to exploit this internal and external congruence, and that piece of folk wisdom demonstrates the epistemic relevance not only of vision's reliability but also of the error-detecting and error-correcting procedures of our broader vision-involving practices. Weinberg (2007, 325)

I have described above some ways that IBE can lead us astray. In this section, I want to defend the claim that despite its fallibility, IBE is *hopeful*. Then I will suggest that hopefulness also helps us mitigate the unwanted consequence of permissivism. Permissivism is problematic because, given that S and S' are both rational and yet arrive at difference verdicts about H, we have no reason to think that S is more likely than S' to have gotten things right. This should undermine S's confidence in her assessment of the evidence. I will suggest that this is right, but it is not a bad thing. We should be hopeful that despite the possibility of permissible disagreement even between ideally rational inquirers, further application of IBE will yield intersubjective agreement. Before discussing this hard case, though, it will be helpful to think about how IBE can be self-correcting for non-ideal inquirers.

One of the features of visual perception that makes it hopeful is that it is to a large-degree self-correcting—when its outputs are potentially misleading, we are often perceptually justified in believing that they are.

> In addition to vision's basic reliability and checkability, it has the very useful property of having outputs that carry their own information as to whether the conditions for its use are sub-optimal. We have a very serviceable sense of the range of conditions and applications in which vision is, and is not, reliable. We know to put less confidence in perceptual judgments requiring a particularly exacting degree of precision, or conducted in poor illumination. Our ability to adjust our confidence is partly the product of cultural inheritance and of our lives—trial and error, but importantly the character of visual experience itself is sensitive to these issues. . . So here we have another important mitigating factor for vision's fallibility: our practices with vision incorporate a capacity to detect under which circumstances it will be especially fallible. The fallibility of vision is much less threatening to the extent that one can anticipate where it will not merely be fallible but will, in fact, be prone to failure. When we find ourselves in such circumstances, we can take steps to keep ourselves from being suckered. (Weinberg, 2007, 326)

IBE can similarly provide information about the circumstances under which it is unreliable. I often know to put less confidence in my inferences from the evidence when I have an emotional stake in the matter. Moreover, I know to do this because when I am emotionally invested in whether or not P, it is often an intellectually satisfying explanation of why I regard P as likely or unlikely that I want or fear that P.

Suppose, for example, that S finds H to be a more satisfying explanation of EK than H', and so has $cr_{\emptyset}(H|EK) > cr_{\emptyset}(H'|EK)$. When this happens, under normal circumstances, the fact that S finds H to be a more satisfying explanation of EK than H' is something that S will also be aware of; that is, her total evidence will come to include this fact (call it Q). Then it may well become a salient question what explains Q, given EK. Let us suppose that in fact the most satisfying explanation of QEK is B, the hypothesis that S is biased in favor of H because she prefers that H is true—or at least that this explanation is at least as satisfying as the explanation that S is rationally responding to the evidence. It may well be true that the fact that S prefers for H to be true is not supported just by IBE, but is independently accessible to S, and may also be added to her evidence. In that case, H will likely no longer be a satisfying explanation of $BQEK$.

Of course, such self-awareness is by no means a trivial achievement, and one may well not know or believe that one is biased when one is. So while IBE *can* be self-correcting, it might be objected that there's no reason to think it is hopeful, because the hypothesis that we are biased or otherwise making an mistake will not necessarily be a lovely explanation, for S, of why S finds H to be a satisfying explanation of EK even when it is the correct explanation. I agree that if a high degree of self-awareness was a necessary condition for IBE to be hopeful, this would not inspire much hope. Fortunately, I do not think that it is necessary.

For, if we introduce additional inquirers into the story, we will get a better source of evidence about when we've made a mistake. As Christensen (2007b) argues, when I learn that an epistemic peer (someone who I reasonably regard as just as well positioned to assess the evidence as I) disagrees with me, I ought to become less confident that I am correct, and more confident that I have made a mistake. And I should do so to the extent that I have no independently satisfying explanation of why my peer has made a mistake than that I have (independently, that is, of my prior degree of conviction in my own judgment).

> These cases, then, suggest the following (admittedly rough) principles for assessing, and reacting to, explanations for my disagreement with an epistemic peer: (1) I should assess explanations for the disagreement in a way that is independent of my reasoning on the matter under dispute, and (2) to the extent that this sort of assessment provides reason for me to think that the explanation in terms of my own error is as good as that in terms of my friend's error, I should move my belief toward my friend's. (Christensen, 2007b, 199)

First, note the appeal to explanatory considerations. What one should do in the face of disagreement with an epistemic peer is determined by what the best explanation of that disagreement is. So, plausibly, the claim that one should compromise by adjusting one's credence in the face of peer disagreement is not an independently motivated norm on reasoning, but a case of inference to the best explanation.

So, for example, I might fail to receive an expected phone call, and become angry, on the hypothesis that the person who did not call deliberately ignored her promise to do so. Perhaps, however, I will tell you about it, and you will disagree with my assessment of the evidence—you find it more likely that the non-caller's phone battery died, or that she was otherwise prevented by circumstances from making the expected call. Supposing you know the person in question and other relevant features of the situation as well as I do, this ought to make me rethink my judgment. In particular, I should look for an explanation of *why* you disagree with me about the evidence. In this case, I will see that my emotional investment has clouded my judgment, and that my explanation of the evidence seemed satisfying only because, in my disappointment, I failed to consider other alternative explanations. That is, the fact of your disagreement provides additional evidence that requires explanation. It is *unexpected* that you disagree with me, given that you're an epistemic peer, or at least it should be if I am convinced of my assessment of the evidence. And as I claimed in the previous section, rational persons are curious about unexpected phenomena. Even when no obvious explanation of my error is forthcoming, unless I have an independent explanation of why you would have reached an incorrect verdict, I ought to regard the hypothesis that I have erred as just as satisfying as the hypothesis that you have. But of course, an increase in my credence in the hypothesis that I have made an error should force a revision in credences based on that possibly erroneous reasoning.

If we apply this to the worrying case of disagreement in judgments of intellectual satisfaction between ideally rational peers, we can see cause for hope. For suppose S and S' are both ideally rational inquirers. Together, they examine the evidence E, explicitly make additional assumptions K, and consider two competing explanations, H and H'. S finds H more satisfying than H', while S' does not. This disagreement provides our subjects with additional evidence that requires explanation—again call this additional evidence Q. Given that they treat each other as peers (and especially if it is part of their assumptions that they are both ideally rational!), S and S' ought to find their disagreement highly unexpected.

Given that they are ideally rational, we can assume no explanation in terms of an inappropriate bias will be forthcoming, and we've stipulated that they are carefully assessing precisely what would explain EK, so we cannot assume that they are making the mistake of relying on false assumptions.

IBE, on my view, requires that S have $cr_{\emptyset}(H|EK) > cr_{\emptyset}(H'|EK)$, while S' permissibly does not. So this is a permissive case—given *only* the evidence EK, there is not a unique, rationally permissible credence that one should have in H. But this will not necessarily lead to any unwanted skeptical consequences. For while I allow that there may be more than one permissible prior probability distribution, the actual *posterior* credence that S and S' should have in H is the credence that would result from conditionalizing their priors on their total evidence, and if Christensen (2007b) is right about how we should respond to disagreement, this will require that they both adjust their credence towards one another's on learning that they disagree. That is, it is consistent with PERMISSIVISM that S and S' are both epistemically obliged to have $cr_{\emptyset}(H|EKQ)$ equal to the mean of their credences in H conditional on EK alone (or whatever other compromise strategy turns out to be rational in the face of such disagreement).

In this case, EK just is not sufficient evidence to settle the question of what credence in H is reasonable. But it is possible to live with cases like this *if*, in every such case, there will be additional evidence available that will determine a unique permissible posterior credence in H. If so, then permissive cases will just point to the need for yet more evidence if we want to settle the question whether H. In cases where I find H intellectually satisfying but others do not, I should not trust that *my* intellectual satisfaction is a guide to the truth of H without further evidence—that would be hubristic. But this does not show that I should not obey IBE. For doing so, and treating intellectual satisfaction as a source of justification, is *hopeful*. By engaging in collaborative inquiry and checking myself against my epistemic peers, I will be able to detect cases in which, through IBE, I arrive at a non-uniquely permissible posterior credence. And so while OBJECTIVE CONDITIONALIZATION entails that if $cr_{\emptyset}(H|EK) = x$ is permissible, then so is $cr_{EK}(H) = x$, this will not be a stable resting place—evidence that another has equally permissible but different conditional prior credence on EK will cause us both to rationally revise our credence in H. In such cases, the best explanation of why I have $cr_{EK}(H) = x$ will *not* be that this credence is what the evidence uniquely supports.

I anticipate several objections to this view. First, that it is circular—I have claimed that it is possible for ideally rational agents to differ in which explanations they find intellectually satisfying, and that intellectual satisfaction determines, via IBE, which prior credences it is permissible to have. I have also claimed that one is required not to find "I am right and you're wrong" to be an intellectually satisfying explanation of evidence of peer disagreement, without some independently satisfying explanation of *why* it is more likely that you've made a mistake. It might seem that I am inappropriately assuming that all ideally rational inquirers *will not* find it more satisfying to just assume that they are right. But I think the assumption that they should not can be independently motivated. If we disagree, then we need to consider the possibility that you've erred and the possibility that I have erred. If we have only symmetrical evidence that favors either hypothesis, as we do if the only evidence is the fact of our disagreement, we should give them equal credence. That is, this is a case in which the Principle of Indifference would seem to constrain our credences. If you disagree, and you claim to find it more intellectually satisfying to suppose that your priors are correct and mine are not, I *will* have an independent explanation of why I am right and you're wrong in this subsequent disagreement, and so I need give your view no weight (I will also take it as evidence that you're not ideally rational).

Another objection is that my explanation of why I should not stick to my guns in the face of disagreement depends on giving credence to the hypothesis that I have made a mistake. But I have also claimed that the credence I have prior to learning of the disagreement is permissible. So where is the mistake? Why is not it a good explanation of the disagreement that neither party has made a mistake, as both have permissible credences? On the one hand, I think that we should say this! Neither party has made a mistake in terms of OC, or IBE. But again, this is because in this sort of case, the evidence does not uniquely determine a rational credence. If we suppose this can happen, we need not also suppose that it is desirable. We might rationally expect that the evidence will uniquely determine a rational credence, at least in the long run (we might have to gather a lot of evidence). Thus we can distinguish between what the evidence will eventually support and what is an intellectually satisfying explanation of the evidence we have right now. If we assume that, given enough evidence, there will be a unique rational credence in H supported by the evidence, then we can distinguish a different sense of mistake, which is the sense in question. Suppose that for all permissible prior probability functions p, there is some evidence E^* that includes my current evidence and background knowledge EK that is sufficient to determine that $p(H|E^*) = x$.[21] Then if I now how have $cr_{EK}(H) \neq x$, I ought to regard my credence in H as possibly mistaken, in the sense that I could have a better-informed credence. Not that this is something I could easily know. But surely it is possible. And we might take it that disagreement is evidence that my credence is not the credence that, in the long run, I ought to have. After all, on the assumption that there is such a credence, at least one of us does not have the correct long-run, best-informed credence. So I should be open to adjusting my credence in the light of new evidence. So I should not stick to my guns just because my current credence was determined in a permissible fashion, and is justified for me, now. If it is possible to do better, I should.

This kind of mistake is not so unfamiliar—we should seek consensus, and try to gather sufficient evidence to ensure it. It would not be epistemically valorous to start with an attitude of suspension of judgment about, say, the proposition that human-generated carbon emissions contribute to global warming when that was still a reasonable attitude even for the best-informed scientists, and then shut one's eyes and ears to the evidence that we've discovered since then, so as to avoid having to revise one's opinion. Similarly, suppose that I was around in the early days of research into this question, examined the evidence, and instead of suspending judgment, found that the explanation in terms of human-generated carbon emissions was intellectually satisfying, and so gave it high credence and believed it. If I then found out that many well-informed climate scientists gave lower credence to the hypothesis, and thought further research was needed, it would be irrational of me to maintain my high credence and refuse to look for better evidence to support my view—and this is true even if I happen to have precisely the high credence that everyone else will have once all the evidence is in.[22] Things are different now, as we have more evidence, and at

[21] For this claim to be non-trivial, E^* must not logically imply or contradict H.

[22] This is not to say that I may not nevertheless regard it as reasonable to act on the assumption that the hypothesis is true, given the high costs of inaction on that assumption and the comparatively low costs of action if it turns out to be false.

present we can easily explain away the disagreement of climate change deniers by looking into who pays their salaries and/or contributes to their political campaigns.

Of course, it is also possible to explain this sort of case by saying that in the early days of research into this question, there was still a unique permissible credence supported by the available evidence, and any divergence from that credence has to be explicable either because the scientists in question were looking at a subset of the evidence (or a superset including some misleading evidence), or they were incorrectly assessing the evidence. I am open to this view; it may be that some are too skeptical and others not skeptical enough, in their evaluations of the evidence. However, I think it is also a *prima facie* reasonable conclusion to draw from this sort of case that differing degrees of skepticism are permissible, and so are varying degrees of confidence in the hypothesis. It is then incumbent on us to gather enough evidence that these differing degrees of skepticism wash out, and consensus emerges. Only then can we suppose that the evidence we've now gathered really does support a unique permissible credence in the hypothesis.

So I claim that PERMISSIVISM is acceptable as long as we also endorse the following:

> SYMMETRY. If S and S' are ideally rational agents and have priors p and q respectively such that $p(H|EK) \neq q(H|EK)$, p and q are both epistemically permissible priors.

> COMPROMISE. If $p_1, p_2, \ldots, p_n$ are all permissible priors, then if P is the proposition that says exactly which priors are permissible, then if $cr_\emptyset$ is a permissible prior, $cr_\emptyset(H|EP) = \frac{cr_\emptyset(H|E)+\sum_{i=1}^{n} p_i(H|E)}{n+1}$ (or some other suitable compromise strategy).

> CONSENSUS. If S and S' are ideally rational agents and have priors p and q respectively such that $p(H|E) \neq q(H|E)$, then $\exists E^*$ such that E is a part of E^*, and E^* is evidence that S and S' could share, but $p(H|E^*K) = q(H|E^*K)$.

If we accept SYMMETRY and COMPROMISE, then there is good reason to think CONSENSUS is true. However, one might now worry that the intersubjective agreement arrived at through COMPROMISE has no relation to the truth. Why should I think that I am more likely to get things right in the long run through COMPROMISE than by sticking to my guns when I have got a permissible credence? In fact, is not COMPROMISE likely to force suspension of judgment on me? Even if we suppose that CONSENSUS is possible, are not we likeliest to agree on a middling credence that mandates suspension of judgment, if large degrees of disagreement are permissible?

Here I appeal again to §3.3.2 and CURIOSITY. There are limits on what ideally rational inquirers find intellectually satisfying. At the very least, they do not find unstable explanations of surprising phenomena satisfying. So we will not, for example, have to compromise with climate change skeptics who posit *no* explanation for increasing global temperatures and simply assert that no explanation is needed. Such individuals would presumably take any amount of evidence for the explanation in terms of human-generated carbon emissions to be irrelevant, and so their credence in that hypothesis will remain low as the evidence piles up, dragging down our compromise credence if we are forced to compromise with them. But such an individual is clearly not ideally rational; she fails to have an adequate degree of curiosity.

To the extent that there is a large degree of divergence in actual (rather than permissible) credences, we can *hopefully* use IBE to detect flaws in many of them; to rule them out as impermissible and give them no weight in our compromise. Thus we can *hopefully* reach a consensus that will support belief or disbelief in the hypotheses we seek to confirm or deny. I agree, however, that it is possible that we will not or cannot. Ultimately I have no answer to this objection except the claim that IBE is hopeful, and we should trust hopeful sources of justification even in the absence of any guarantee of their reliability. I do think that the best explanation of why we *do* trust intellectually satisfying explanations and seem to have made good epistemic progress by doing so is that inferring the most intellectually satisfying explanation (and seeking intersubjective consensus and compromise) is a good way to arrive at the truth in the long run. But of course, that will not convince anyone who does not already agree.

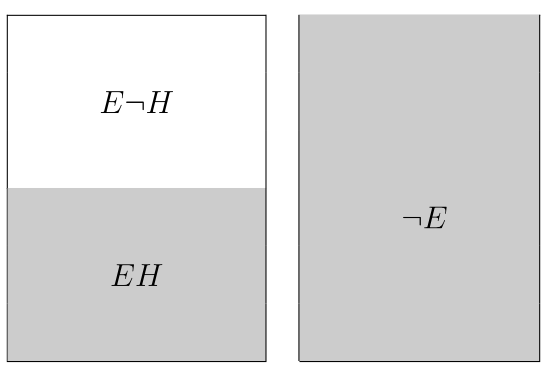
3.4 RATIONALIST LIBERALISM AND IBE

I claimed that I was going to argue that when S finds H to be an intellectually satisfying explanation of E, S is sometimes immediately justified in believing $E \supset H$. I have not mentioned $E \supset H$ in my account of IBE, so far. Instead, I have claimed that if S is ideally rational and finds H to be a more satisfying explanation of E than H', given additional assumptions K, S should have priors such that $cr_{\emptyset}(H|EK) > cr_{\emptyset}(H'|EK)$.

As shown in §2.2.2, if p is a probability function, then $p(H|E) = 1 - \frac{1-p(E \supset H)}{p(E)}$. So it is easy to show that if $p(H|E) > p(H'|E)$, then $p(E \supset H) > p(E \supset H')$. Therefore, IBE entails that an ideally rational subject S who finds H to be a more intellectually satisfying than H' as an explanation of E, given K, should have priors such that $cr_{\emptyset}(E \supset H|K) > cr_{\emptyset}(E \supset H'|K)$.

But it is also true that an increase in $p(E \supset H)$ while holding $p(E)$ fixed will force an increase in her credence in H conditional on E. So if S gains justification to believe $E \supset H$ when she finds H to be a satisfying explanation of E, this will be a way of satisfying the constraint that IBE imposes. Another way to think about the probabilistic relationships here is as follows. $\neg E$ entails $E \supset H$, so $p(E \supset H) \geq p(\neg E)$ for all probability functions p. In fact, the precise amount by which $p(E \supset H)$ exceed $p(\neg E)$ is just that portion of the region of probability space occupied by E that overlaps the region occupied by H, i.e, $p(E \supset H) = p(\neg E) + p(EH)$. Graphically, $p(E \supset H)$ is the shaded region of Figure 4 below. Now, if we hold $p(E)$ fixed, we cannot increase the width of the cells of the table, so any increase to $p(E \supset H)$ must obviously increase the portion of $p(E \supset H)$ that lies inside the region representing $p(E)$. But this means a proportional increase to $p(H|E)$, i.e., in the extent to which E confirms H.

Figure 4: PROBABILITY OF $E \supset H$



Now, if we are considering what would hypothetically be a good explanation of some evidence E, ordinarily this will not cause us to shift our prior credence in E—we are not wondering how likely E is, just what a good explanation of E would be if we were to learn

E. So suppose I have prior credence distribution such that $cr_K(H|E) = cr_K(H'|E)$, and I consider what would be a satisfying explanation of E. I find H to be more satisfying than H'. If I adjust my hypothetical priors $cr_\emptyset$ so that $cr_\emptyset(E \supset H|K) > cr_K(E \supset H)$, then holding my credence in E, H, and H' fixed, this means increasing $cr_\emptyset(HE|K)$—the subregion of E that overlaps with H. But this is precisely the same as increasing my credence in H conditional on EK. So with my new priors, I will have $cr_\emptyset(H|EK) > cr_\emptyset(H'|EK)$.

Given my strong objective Bayesian account of IBE, if I find H to be an intellectually satisfying explanation of E given K (assuming I am sufficiently curious and unbiased), it is not permissible for me to have the priors that I started with, and it is permissible for me to adjust my priors so as to have higher credence in $E \supset H$, conditional on EK. But if I did not find H to be a satisfying explanation of E given K, possibly my original priors would have been permissible. So it is the fact that I find this explanation satisfying that obliges me to give higher credence to $E \supset H$. And I may be obliged to have this high credence whether or not I give high credence to E or H. So I am justified in giving higher credence to $E \supset H$ because I find H to be a satisfying explanation of E, independently of my credence in E or H.

My take on IBE supports a rationalist liberal view about inferential justification. For suppose I learn E with background knowledge K, and I ought to have posterior credence such that $cr_{EK}(H) = x$, where x is relatively high—E is good evidence for H, given K. Then there is a permissible prior probability function p such that $p(H|EK) = x$, with $p(E \supset H|K) \geq p(E \supset H|EK) \geq x$. Either my actual priors cr_K are equal to $p(\cdot|K)$, or not. If they are, then I was already justified in having a high credence in $E \supset H$, conditionally on E. If they are not, then I *now* ought to have those priors. But then there has to be something that explains why it is rational for me to adjust my priors in this way. Here, IBE can do the job of explaining why my actual priors are impermissible. But if it is IBE that justifies adjusting my credence in this non-conditionalizing way, then I will be justified in having priors such that $p(E \supset H|K) \geq x$ because H is an intellectually satisfying explanation of E, given K. But this means that I am justified in having a high credence in $E \supset H$ on the basis of K alone, independently of E or H. So I am justified in having a high credence in $E \supset H$ prior to and/or independently of E and H. After all, I could have made this adjustment prior to learning E, just by considering that H would be a satisfying explanation of E, were I to learn E. So my view is compatible with RATIONALISM.

I think we can also preserve the spirit of LIBERALISM given my view. For suppose I already permissibly had $cr_K(H|E) = p(H|EK)$. Then I need not make any adjustment to my credence in $E \supset H$, and I need not believe it, or even have ever considered it, in order to arrive at a high posterior credence in H by conditionalization. Or if I learn E and consider for the first time what would be a good explanation of it, and I decide that H provides the most satisfying explanation, I need not, as in a heuristic view of IBE, take this as *evidence* that functions p with $p(E \supset H|K) < x$ are impermissible; I might just (rationally) adjust my priors and conclude H on the basis of E. As I have complained before, it is a little bit hard to know exactly what it means to say that one's justification for H "includes" justification to believe some other proposition. But at the very least, I do not think that learning that H is a satisfying explanation of E need make it rational to believe in a counterfactual supporting relationship between E and H, or anything like that (as Fumerton would) in order to make it rational to infer H from E. Nevertheless, as a byproduct of adopting priors that license the inference, one will reasonably have a high credence in $E \supset H$, conditionally on K.

3.5 CONCLUSION

I do not claim to have answered every possible skeptical worry about inferential justification. But it is a requirement that: when E confirms H, one must have a relatively high prior credence in $E \supset H$. I have argued that this high prior credence must be justified independently of E; dogmatism about inferential justification is incompatible with most probabilistic models of inferential justification. I have also offered an account of inference to the best explanation that explains why we often should have high prior credence in $E \supset H$; when H is the most intellectually satisfying explanation of E, we are epistemically obliged to give H higher prior credence conditional on E than alternative explanatory hypotheses. Given some plausible assumptions about how IBE should apply in cases of disagreement and about the epistemic impermissibility of being incurious, we have reason to think that we can reach consensus on a single epistemically permissible credence. We should be hopeful that the theoretical consensus credence that will be permitted by sufficiently weighty evidence will be decisive, in the sense of permitting belief or disbelief in the hypothesis in question.

BIBLIOGRAPHY

Arntzenius, Frank. 2003. "Some Problems for Conditionalization and Reflection." *The Journal of Philosophy* 100:356–370.

Burge, Tyler. 1993. "Content Preservation." *Philosophical Review* 103:457–88.

Casullo, Albert. 1988. "Revisability, Reliabiliism, and A Priori Knowledge." *Philosophy and Phenomenological Research* 49:187–213.

Christensen, David. 2004. *Putting Logic in Its Place*. Oxford University Press.

—. 2007a. "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals." In *Oxford Studies in Epistemology*, volume 2, 3–31. Oxford University Press.

—. 2007b. "Epistemology of Disagreement: The Good News." *Philosophical Review* 116:187–217.

—. 2010. "Higher-Order Evidence." *Philosophy and Phenomenological Research* 71:185–215.

Cohen, Stewart. 1998. "Contextualist Solutions to Epistemological Problems: Skepticism, Gettier, and the Lottery." *Australasian Journal of Philosophy* 76:289–306.

D. Kahneman, P. Slovic and Tversky, A. (eds.). 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.

DeRose, Keith. 1992. "Contextualism and Knowledge Attributions." *Philosophy and Phenomenological Research* 52:913–929.

—. 1995. "Solving the Skeptical Problem." *Philosophical Review* 104:1–52.

Dougherty, Trent and Rysiew, Patrick. 2009. "Fallibilism, Epistemic Possibility, and Concessive Knowledge Attributions." *Philosophy and Phenomenological Research* 78:123–132.

Dretske, Fred. 1970. "Epistemic Operators." *The Journal of Philosophy* 67:1007–1023.

—. 2005. "The Case Against Closure." In Ernest Sosa and Matthias Steup (eds.), *Contemporary Debates in Epistemology*, 13–25. Oxford: Blackwell.

Elga, Adam. 2000. “Self-Locating Belief and the Sleeping Beauty Problem.” *Analysis* 60:143–147.

—. 2010. “Subjective Probabilities Should Be Sharp.” *Philosophers' Imprint* 10.

Fantl, Jeremy and McGrath, Matthew. 2002. “Evidence, Pragmatics, and Justification.” *The Philosophical Review* 111:67–94.

—. 2007. “On Pragmatic Encroachment in Epistemology.” *Philosophy and Phenomenological Research* 75:558–589.

—. 2009. “Advice for Fallibilists: Put Knowledge to Work.” *Philosophical Studies* 142:55–66.

Feldman, Richard and Conee, Earl. 2004. “Evidentialism.” In *Evidentialism*, 83–107. Oxford University Press.

Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Rowman & Littlefield.

Gettier, Edmund. 1963. “Is Justified True Belief Knowledge?” *Analysis* 23:121–123.

Hájek, Alan. 1998. “Agnosticism Meets Bayesianism.” *Analysis* 58:199–206.

—. 2008a. “Arguments for—or Against—Probabilism?” *British Journal for the Philosophy of Science* 59:793–819.

—. 2008b. “Dutch Book Arguments.” In Prasanta Pattanaik Paul Anand and Clemens Puppe (eds.), *The Oxford Handbook of Rational and Social Choice*. Oxford University Press.

Hájek, Alan and Hartmann, Stephen. 2010. “Bayesian Epistemology.” In Jonathan Dancy, Ernest Sosa, and Matthias Steup (eds.), *A Companion to Epistemology*, 93–106. Blackwell, 2nd edition.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Clarendon.

—. 2005. “The Case for Closure.” In Ernest Sosa and Matthias Steup (eds.), *Contemporary Debates in Epistemology*, 26–42. Blackwell.

Hawthorne, John and Lasonen-Aarnio, Maria. 2009. “Knowledge and Objective Chance.” In Patrick Greenough and Duncan Pritchard (eds.), *Williamson on Knowledge*, 92–108. Oxford University Press.

Jeffrey, Richard C. 1965. *The Logic of Decision*. McGraw-Hill.

—. 1970. “Dracula Meets Wolfman: Acceptance vs. Partial Belief.” In Marshall Swain (ed.), *Induction, Acceptance, and Rational Belief*. D. Reidel.

Joyce, James. 2005. “How Probabilities Reflect Evidence.” *Philosophical Perspectives* 19:153–178.

—. 2011. “A Defense of Imprecise Credences in Inference and Decision Making.” In *Oxford Studies in Epistemology*, volume 4. Oxford University Press.

Keynes, John Maynard. 1921. *A Treatise on Probability*. Macmillan.

Kung, Peter. 2010. “On Having No Reason: Dogmatism and Bayesian Confirmation.” *Synthese* 177:1–17.

Lewis, David. 1996. “Elusive Knowledge.” *Australasian Journal of Philosophy* 74:549–567.

Lipton, Peter. 2004. *Inference to the Best Explanation*. Routledge, 2nd edition.

McGrew, Timothy. 2003. “Confirmation, Heuristics, and Explanatory Reasoning.” *British Journal for the Philosophy of Science* 54:553–567.

Nozick, Robert. 1981. *Philosophical Explanations*. Harvard University Press.

Okasha, Samir. 2000. “Van Fraasen’s Critique of Inference to the Best Explanation.” *Studies in the History and Philosophy of Science* 31:691–710.

Pryor, Jim. 2000. “The Skeptic and the Dogmatist.” *Noûs* 34:517–549.

—. 2004. “What’s Wrong With Moore’s Argument?” *Philosophical Issues* 14:349–378.

Ross, Jacob and Schroeder, Mark. forthcoming. “Belief, Credence, and Pragmatic Encroachment.” *Philosophy and Phenomenological Research* .

Salmon, Wesley. 2001. “Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation.” In G. Hon and S. S. Rakover (eds.), *Explanation: Theoretical Approaches and Applications*. Kluwer.

Silins, Nicholas. 2007. “Basic Justification and the Moorean Response to the Skeptic.” *Oxford Studies in Epistemology* 2:108–137.

Skyrms, Brian. 1980. *Causal Necessity*. Yale University Press.

Sturgeon, Scott. 2008. “Reason and the Grain of Belief.” *Noûs* 42:139–165.

van Fraassen, Bas. 1989. *Laws and Symmetry*. Oxford University Press.

Weatherson, Brian. 2005. “Can We Do Without Pragmatic Encroachment?” *Philosophical Perspectives* 19:417–443.

—. 2007. “The Bayesian and the Dogmatist.” In *Proceedings of the Aristotelian Society*, volume 107, 169–185. Blackwell.

Weinberg, Jonathan. 2007. “How to Challenge Intuitions Empirically Without Risking Skepticism.” *Midwest Studies in Philosophy* 31:319–343.

Weintraub, R. 2001. "The Lottery: A Paradox Regained and Resolved." *Synthese* 129:439–449.

Weisberg, Jonathan. 2009. "Locating IBE in the Bayesian Framework." *Synthese* 167:125–143.

White, Roger. 2005a. "Epistemic Permissiveness." *Philosophical Perspectives* 19:445–459.

—. 2005b. "Explanation as a Guide to Induction." *Philosophers' Imprint* 5.

—. 2006. "Problems for Dogmatism." *Philosophical Studies* 131:525–557.

—. 2010. "Evidential Symmetry and Mushy Credence." *Oxford Studies in Epistemology* 3:161–188.

Willenken, Timothy. 2011. "Moorean Responses to Skepticism: A Defense." *Philosophical Studies* 154:1–25.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford University Press.

—. 2009. "Probability and Danger." *The Amherst Lecture in Philosophy* 4.