Against Explanationist Bayesianism

Lok C. Chan

1 Bayesianism and Explanationism

In response to van Fraassen's argument, many arguments for compatibilism from the explanationist perspective have emerged. In particular, I will focus on Okasha (2000), Lipton (2004), and Weisberg(2009). There are important subtle differences in their views, but these explanationists do share a common strategy in pointing out that van Fraassen's characterization of PIBE is not mandatory, and that PIBE can be construed in a way that renders it not only compatible with, but also, so they argue, complementary to Bayesianism. Okasha's way is to hold that what PIBE does is to facilitate the computation of subjective priors and likelihoods. Lipton suggests that PIBE is a heuristic human beings use to achieve conditionalization is a more efficient way. Both of these variants of compatibilism tries to solve the conflict by focusing on the psychological aspect of Subjective Bayesianism. Weisberg, on the other hand, takes to objective route: he argues that what PIBE does is to characterize some objective *a priori* distribution p. I shall argue, however, that all three attempts are inadequate. I will also use my analysis of these positions as an occasion to explain

abductive volunatarism.

1.1 Probabilism and Bayesianism

For the sake of clarity, we must first decide what being a "Bayesian" entails. It certainly won't do to call anyone who uses Bayes' theorem a Bayesian, since the theorem itself is a logical consequence of the theory of probability, so anyone who accepts the axioms of probability is a "Bayesian" in that sense. We also ought to distinguish Bayesianism in epistemology with Bayesianism in statistics. Even the strictest Bayesian statistician, who might reject all uses of Frequentist methods, could consistently maintain that Bayes' theorem is not the *only* rational way to revise one's belief in light of new evidence.

Since explanationists are responding to van Fraassen's claim that Bayesianism and explanationism is incompatible, I suggest that we ought to retain the language

I will begin by spelling out a foundational thesis that encompasses the types of views relevant to our discussion: *probabilism* is the acceptance of the two statements below:

1. There are *partial* beliefs, which are measurable in terms of degrees.
2. The axioms of probability govern the rationality of partial beliefs.

Strict Bayesianism, as an epistemological thesis, is the position that accepts the tenets of probabilism, plus the thesis that *conditionalization* is only rational way to revise one's belief:

Conditionalization: it is rational to update one's prior degree for belief

H in light of evidence E via the application of Bayes' theorem, which determines the posterior opinion degree of belief $P(H|E)$.

Thus, Strict Bayesianism claims that that a rational agent is one whose belief can be captured by some prior distributions, which would be mechanically revised whenever data is introduced. In our particular context, the term "Bayesian" refers to someone who thinks that not only is this revision justifiable only by conditionalization, but also that rationality mandates it. A person does not have to a Bayesian to think that conditionalization can be just the right thing to do. One, for instance, may think that while conditionalization is a perfectly fine principle, it cannot be the only rule, since there are occasions in which it is not applicable, e.g., a required prior might not be available.

1.2 Explanationism

In its most naive form, inference to the best explanation(IBE) says that we should infer toμ the hypothesis that best *explains* the evidence we have. In our context, explanationism refers to the position that IBE is viable *inductive* rule of inference. More significantly, Harman's version, IBE is conceived to replace enumerative induction, that is, inferring from samples to the population.[1] The strongest, which is not the topic of our discussion can be called *naive inference to the best explanation*:

NIBE: One is entitled to infer from the the fact that a hypothesis H would best explain the evidence E to the conclusion that H is true.

---

[1] Gilbert H. Harman, "The Inference to the Best Explanation," *Philosophical Review* 74, no. 1 (1965): 90.

In this chapter, I will not concern myself with naive explanationism. Instead, my focus is on the explanationists who accept and argue for what I call the *compatibilist* thesis:

> *Compatibilism*: Inference to the best explanation is compatible with some form of Bayesianism.

The compatibilists are almost universally motivated by van Fraassen's argument against a non-compatiblist version of Explanationism. Van Fraassen is well-known for arguing against IBE in its original, non-probabilistic, form. In its most naively powerful form—a view that van Fraassen does ascribe to some philosophers—IBE can be construed as a solution to Hume's problem of induction, which holds that there is no independent justification for extrapolating inductively beyond the evidence we have. IBE gets us out of this problems by giving justificatory force to explanatory virtues, so that the best explanation is the one we *should* accept. Van Fraassen attacks this position relentlessly. One often cited argument of his is that we never pick the best explanation *simplicter*, but the best *out of the explanations available to us.* Van Fraassen argues this is a horrible justification for a belief, since for some reason we might only have horrible explanations available to us, so 'our selection may well be the best of a bad lot.' (143)

Van Fraassen suggests that the strongest recourse available the supporters of IBE is *entrenchment*, which amounts to the repackaging IBE into a rule that works well with Bayesianism. The more plausible way to do this, according to van Fraassen, is to give explanatory virtues a place in the revision of belief in light of new evidence: 'behind the naive rule of IBE there might lie a recipe for adjusting our personal

probabilities, in response to new experience, under the aegis of explanatory success'. Let's call this new rule 'probabilistic inference to the best explanation', which entitles us to raise the probability of the best explanation:

PIBE: One is entitled to infer from the the fact that a hypothesis H would best explain the evidence E to the conclusion that H is *more probable.*

Van Fraassen, however, argues that this cannot do. To begin, if IBE is to be harmonized with Bayesianism, since it must not clash the Bayesian procedure of belief revision, i.e., conditionalization, but van Fraassen insists that this cannot be done without violating the Bayesian standard of rationality. To begin, he suggests that the new IBE-inspired rule must be *ampliative*: it goes beyond what is logically implied by our evidence. In contrast, Bayes' theorem is *explicative*, since the posterior probability is nothing but an arithmetic consequence of conditional probability. This must mean that the PIBE is the rule that confers 'bonus' probability to a belief based on its explanatory virtue. This is where PIBE conflicts with Bayesianism.

A simplified version of his argument is as follows. Suppose we are interested in the bias of a certain coin, θ, which indicates the probability of the coin landing on heads. Suppose we know that there are 3 equally probable hypotheses: (A) $\theta = 0.9$, (B) $\theta = 0.5$, and (C) $\theta = 0.1$. Suppose our evidence gathering process is described as follows: $X_i = 1$ denotes 'the coin has landed on heads on the i toss' and $X_i = 0$ otherwise. Suppose we have toss the coin 4 times, and they all landed on heads. So the evidence E is $\sum_{i=1}^{4} X_i = 4$. The marginal probability for E is:

$$P(E) = P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C) = 0.24$$

Using Bayes' theorem, the posterior probabilities are: $P(A|E) = 0.9129$, $P(B|E) = 0.0869$, and $P(C|E) = 0.0001$. So far so good—4 consecutive heads favors the hypothesis that the coin is bias toward heads, which is what conditionalization is showing us. Where does PIBE comes in, then? Van Fraassen asserts that an argument from PIBE would be as follows: out of the three hypotheses, A best *explains* why we see nothing but heads: it's because it's highly biased. So what PIBE should do is to recommend the redistribution of the probabilities so that $P(A)$ would be even higher. Suppose we raise $P(A)$ to 0.999. This amounts to giving the best explanation a bonus of 0.086 in probability. To accommodate this, we can lower the probabilities of the other hypotheses accordingly. For instance: $P(A) = 0.999$, $P(B) = 0.00086$, and $P(C) = 0.00014$. So we have extrapolated ampliatively beyond what the evidence tells us by using PIBE.

This line of reasoning, however, flies in the face of the Bayesian notion of *coherence*, since it renders one subject to an so-called *Dutchbook*, which is a set of bets that ensures whoever takes these bets a loss. Imagine that we are back at the beginning before we tossed 4 heads. Before tossing the coin for four times, we were offered the following set of bets. Let E again be 'the coin is tossed 4 times and they are all heads' and H be the $X_5 = 1$, that is, 'the fifth toss turns up heads'.

1. $10,000 if E and $\neg H$.
2. $1300 if $\neg E$.
3. $300 if E.

Now, we can calculate the values of these bets based on our prior probabilities:

1. Bet 1: $(10000)\frac{0.8^4(0.2)+0.5^5+0.2^4(0.8)}{3} = 323.16$

2. Bet 2: $(1300)(1 - 0.158) = 988.56$

3. Bet 3: $(300)0.158 = 71.87$

So these bets would be worth \$1383.6 in total. Suppose we bought these bets for exactly that much from a bookie, who then proceeded to toss the coin for 4 times. Either E is true or it is false. Suppose it's false—at least one toss landed on tails. In this case, we would have won bet 2 but lost 1 and 3. We would receive \$1300 but this would still lead to a total loss of $-1383.6 + 1300 = -83.6$.

On the other hand, suppose E—all tosses turned up heads. We would receive \$300 per bet 3, and now bet 1 would depend entirely on the fifth toss. Now, van Fraassen asks, what should our degree of belief for $\neg H$, that the fifth toss will land on tails? Recall that we have used PIBE to give a bonus to the most explanatory hypothesis, A, which effectively has raised the marginal probability of H to 0.9. At this point, bet 1 is now worth $(10000)P(\neg H) = (10000)0.1 = 1000$. Suppose the bookie offers us exactly \$1000 to buy this bet back. We would regard it as fair and accept his offer. In this scenario, we end up with $-1383.6 + 300 + 1000 = -83.6$—we incur exactly the same loss as we would if E were false.

We have been "Dutchbooked", which is a sign of irrationality according to Bayesianism. This happens exactly because we went beyond conditionalization with PIBE. Thus, van Fraassen concludes that compatibilism must be false: Bayesianism and Explanationism are inconsistent with each other.

2   The Compatibilist Argument from Psychology

While van Fraassen never spells out PIBE, his argument does assume a certain formulation of it. In particular, he is relying on the idea that PIBE must give bonus probability to the posterior probability after conditionalization. Many explanationists, then, believes that this is an uncharitable interpretation of PIBE and an alternative formulation can open the door to a viable compatibilism.

One way to accomplish this is to make an argument from human psychology, i.e., how people use explanations to assist probabilistic reasoning. The key to this argument is to suggest that the role of PIBE is to inform an area of probabilistic reasoning about which Bayesianism is less opinionated

To begin, note that strict Bayesianism is an incomplete doctrine. It requires an additional claim about the status of prior distributions: since conditionalization presupposes the existence of a prior distribution for the belief H supported by evidence E, Bayesians need an account for the status of these priors. In particular, it must address how one can determine if a prior degrees is admissible to conditionalization. One view is that priors are *subjective*:

> *Subjective Bayesianism*: priors need not be justified, as long as they are consistent with the axioms of probability.

From an epistemological point of view, Subjective Bayesianism can be considered as a type of anti-foundationalism. The idea is that the subjective view frees us from the burden of justifying our current opinions and holds that all rationality demands from us is to update our priors—whatever they are and however they were formed—by

applying the rule of conditionalization diligently. This is an attractive feature of the subjective view, since it dislodges skepticism's most powerful premise, that rationality demands all our beliefs to be justified. Van Fraassen points out that this is what allows Subjective Bayesians to allow reasonable disagreement without succumbing to relativism. The thought is that with enough evidence, two completely opposite opinions could converge, as long as they strictly conditionalization as a rule of belief revision.

The permissive natural of the subjective prior gives explanationists some leeway to find compatibility. One way to advance this argument is to suggest that explanations have traditionally been associated with the prior and likelihood, and not the posterior.

2.1   Must PIBE Confer 'Bonus' Posterior Probability to the Best Explanation?

Naturally, many explanationists see this forgiving attitude toward priors as a door to an alternative reading for PIBE. Recall that van Fraassen's argument relies on the idea that PIBE give bonus, not to the prior probability, but the posterior, which is mechanically derived using Bayes' theorem. A natural response is that there must be something wrong about van Fraassen's argument, because PIBE does not necessarily have to give bonus probability to the posterior. One way to advance this argument is to suggest that explanations have traditionally been associated with the prior and likelihood, and not the posterior. Gilbert Harman suggests that a good explanation should have a high prior probability and a high likelihood—probability of the evidence

conditional on the explanation.[2] That is, take two hypotheses H_1 and H_2, if H_1 is a better explanation of some evidence E than H_2, then

1. $P(H_1) > P(H_2)$ and,
2. $P(E|H_1) > P(E|H_2)$.

Note that the high probabilities are taken as *necessary* conditions of a good explanation. This then opens the door to the idea that, if we start by knowing that a hypothesis is the best explanation we have, we can infer that it has higher prior and likelihood than alternative hypotheses. Samir Okasha argues that not only this characterization "captures the phenomenology of inferring to the best explanation much better than van Fraassen's account", but it also shows that PIBE is "perfectly consistent with Bayesian principles".[3] His idea is that in actual reasoning, explanatory reasoning can help an agent to carry out Bayesian calculations on the fly, without giving a bonus to the best explanation in a way that makes one vulnerable to Dutchbook arguments. To demonstrate what he means, he asks us to imagine the following scenario: a doctor is trying to diagnose if a child has H_M: pulled a muscle or H_L: torn a ligament. Okasha suggests that her reasoning can be a mix of PIBE and Bayesian conditionalization as follows:

> Suppose we ask the doctor to justify her reasoning. She answers: 'firstly,
> pre-adolescent children very rarely pull muscles, but often tear ligaments.
> Secondly, the symptoms, though compatible with either diagnosis, are
> exactly what we would expect if the child has torn a ligament, though not

---

[2]Gilbert Harman, "Knowledge, Inference, and Explanation," *American Philosophical Quarterly* 5, no. 3 (1968): 169.

[3]Samir Okasha, "Van Fraassen's Critique of Inference to the Best Explanation," *Studies in History and Philosophy of Science Part A* 31, no. 4 (2000): 703.

if he has pulled a muscle. Therefore the second hypothesis is preferable.'[4]

Okasha's suggestion is that the doctor has used to explanatory virtues to arrive at the values for the prior probabilities for the child tearing a ligament or pulling a muscle—call them, respectively, $P(H_L)$ and $P(H_M)$—and the probability of evidence, i.e., the symptoms, given each hypothesis, that is, $P(E_S|H_L)$ and $P(E_S|H_M)$. He argument is that since in this case what PIBE does is not to confer any bonus probabilities on the doctor's degrees of belief after conditionalization, there is no threat of being Dutchbooked. Thus, "the conflict between IBE and Bayesianism alleged by van Fraassen depends entirely on an idiosyncratic way of representing IBE in probabilistic terms."[5] His argument, however, is problematic, and does not show what the conclusion he states.

Do we have a genuine case of Bayesian PIBE? It is clear that Okasha intends to use this example to demonstrate how IBE can be used along side of conditionalization, but what is invovled in his example is something quite different. He says that the doctor "has used explanatory considerations as an aid for calculating the priors and likelihoods needed to apply Bayes's theorem itself."[6] Let us recall that the conclusion of a conditionalization is a posterior degree of belief based on the prior distribution and the evidence as expressed by the likelihood distribution. It has to be a numerical value.[7] In other words, conditionalization require the actual values of $P(H_L|E_S)$ and $P(H_M|E_S)$.

But this is not what the doctor does in both the story and Okasha's analysis. By

---

[4]Okasha, 703.

[5]Okasha, 703.

[6]Okasha, 703.

[7]Or an interval, but we can put this detail aside.

concluding which hypothesis is more preferable, what the doctor has inferred is a rough estimate of this inequality, $P(H_L|E_S) > P(H_M|E_S)$. While this inequality might only look superficially different than the two posterior probabilities, the computations involved are quite different, and has important epistemological implication. The important point is that deciding between two competing hypotheses is a different inferential step than conditionalization. To begin, the actual posteriors would require the following computations:

$$P(H_L|E) = \frac{P(H_L)P(E|H_L)}{P(E_S)} \qquad P(H_M|E_S) = \frac{P(H_M)P(E|H_M)}{P(E_S)}$$

Note that E_S, the probability of the child having the symptoms, is required for the calculations of both posterior. To determine this probability, however, we need to know the sum of the probabilities of E_S jointly occurring with each possible hypothesis H_i, assuming that we are dealing with a discrete distribution:

$$\sum_{i=1}^{k} P(H_i)P(E_S|H_i)$$

That is, we need to know the prior probabilities for the every possible disease the child might have, and the probability of her symptom occurring given this disease, assuming the hypotheses are independent. In other words, what we need is the prior probability of the symptoms given the child has neither a torn ligament nor a pulled muscle—this is what Abner Shimony calls the "the catchall", that is:

$$P(\neg H_L \wedge \neg H_M)P(E_S|\neg H_L \wedge \neg H_M)$$

12

For any problem without a well defined model with a predefined space of hypotheses, the catchall can be difficult or even intractable. In this case of making a diagnosis based on symptoms, conditionalization requires the doctor the probability of each possible hypothesis H_i and the conditional probability of the symptom E_S on that hypothesis H_i beyond the two hypotheses mentioned by the doctor, unless, of course, all other hypotheses have a zero prior. However, is this a reasonable representation of how people think when they consider two competing hypotheses? I am inclined to say no—just because an agent singles out two hypotheses as being the most worthy of examination, it does not mean all other hypotheses have zero probability. This is not to say that doctors never use probability in diagnostic problems, but that Okasha's appealing to phenomenology is unfounded. As far as I am concerned, nothing in my first person experience ever suggests that I have carried out conditionalization in my head, with or without the assistance of PIBE.

Perhaps Okasha believes that explanation can only inform the prior and likelihood of hypothesis in question, and has nothing to do with the denominator of Bayes' theorem and the catchall. Still, Okasha's defense appears to me rests on a fundamental misunderstanding of a thesis about prior distributions shared by van Fraassen and Subjective Bayesians. Neither is committed to idea that priors need to be justified for it to be admissible for conditionalization. Once this is pointed out, the reason for van Fraassen's "idiosyncratic" representation is clear as day—he has no problem with PIBE involved with informing the choice of the prior, as long is it is not invoked after one has committed into conditionlizing as a rule.

Okasha seems to be aware of this criticism.[8] His response is that even though his

---

[8]Okasha, 705–6.

version of PIBE only pertains to how priors and likelihoods are to be arrived, it still figures into the updating process of Bayesianism since the priors and likelihoods are crucial components of conditionalization.

This response, however, is tone-deaf to the iterative nature of Bayesian updating. What the posterior now will be the prior for the next iteration. Okasha's characterization of Bayesian reasoning completely sidesteps this crucial procedure, since the only case he considers is one where the prior is yet to be informed by the evidence, and this is when the prior could be more or less chosen arbitrarily by the agent without being worried about being Dutchbooked. In other words, during the doctor's first encounter of the situation, she is free to adopt a prior based on her expertise and experience. At this point, she can use PIBE to her heart's content within the bounds of logic and probability. For instance, if H_L and H_M are the only two mutually exclusive hypotheses, then the sum of their probabilities cannot be more than 1, but other than that she is under no threat of being Dutchbooked until she revises her belief via conditionalization. If all she does is to compute her posterior probability based on her PIBE-inspired priors and likelihoods, then Okasha is correct in suggesting that she will not be Dutchbooked, since her explanationist musings are contained in the subjective initial prior. But, again, this was never van Fraassen's problem with PIBE.

In other words, a crucial problem never addressed by Okasha is the following: suppose the doctor runs some diagnostic on the child to test her hypotheses, e.g., ordering an X-ray, to gather further evidence, E_x. Now the doctor has to revise her degrees of belief on the two hypotheses based on new evidence. Now, the crucial question is: at this stage, could she revise her no longer ignorant priors beyond what

14

is implied by E_x and Bayes' theorem?

Based on Okasha's idea that PIBE just naturally figures into the agent's phenomenology when making probabilistic judgments, it would appear that the doctor would again reconsider the priors and the likelihoods based on their explanatory virtues. So suppose the answer to the question posed in the last paragraph is *yes*—the doctor should adjust the new priors based on their explanatory virtues beyond what conditionalization tells us, she would be essentially giving bonus to the posterior probability in a way that makes herself liable to be Dutchbooked as van Fraassen argued. If the answer is *no*, then I don't see how Okasha has engaged in van Fraassen's argument in all, since after all this means that PIBE has no place in the revision of belief in light of evidence. All it does is to influence a prior before undergoing *any* conditionalization, which is already unrestricted for Subjective Bayesians. So in either case Okasha's account is inadequate.

So, to answer the question: must PIBE confer bonus probability to hypothesis? Not necessarily— one could conceive, as Okasha does, PIBE as being parasitic to priors and/or likelihoods, but I have argued that Okasha's position has been hamstrung at the beginning by relying on the appeal to our intuition or phenomenology on how we reason. X A natural response is the thought that instead of our intuition, one could get support from cognitive science and psychology for the explanationist cause. This is the approach of Peter Lipton, to whom we shall now turn.

2.2 Is PIBE More Psychologically Realistic?

In broad strokes, Lipton's argument is that explanation serves an approximate realization Bayesian of conditionalization. Lipton's argument shares a similar structure as Okasha's, in that they both are suggesting that in day to day reasoning, IBE-like reasoning is how people think about probability, so Bayesian should want compatibilism as much as explanationists do. Lipton's argument, however, is more empirically informed, since it appeals to the classic literature in behavioral economics done by Kahneman and Tversky. Lipton's aim is also more substantive: not only he claims that Bayesianism and PIBE are compatible, but PIBE in fact complements Bayesianism by providing a psychologically realistic picture of people can approximate Bayesian reasoning in real life.

Lipton's suggestion, then, is more ambitious and explicit than Okasha's: Okasha, for the most part, simply gestures toward the idea that explanation-based inferences describes how people could come with priors and likelihoods. Lipton's thesis, on the other hand, is much emphatic: he thinks that the behavioral economics and psychology has shown that people use IBE to compute posterior probabilities. Lipton recognizes that this thesis is clashing directly with van Fraassen's argument:

> . . . .it may sound as though explanatory consideration are somehow to modify Bayesian formula, say by by giving some hypothesis a poster 'bonus' beyond what Bayes's theorem itself would grant. . . This is where the spectre of dutch book irrationality appears. . . but it is not what I propose. Instead I want to suggest some of the ways in which explanatory considerations may be our way of running or realizing the mechanism

of Bayesian conditionalization—the movement from prior to posterior probability—and our way of handling certain aspects of inference that conditionalizing does not address.[9]

How does that work? The first step of Lipton's argument is that behavioral economics has shown that people are in generally bad at thinking probabilistically, and often draw inferences that violate the axioms of probability. The second step is to claim that these findings support Lipton's theory that people often use IBE to reason about probabilistic problem, even though the result is not always optimal. The last step is to conclude from the previous steps that Bayesianism fares poorly against explanationism as a descriptive theory about human being, so that an irenic approach benefits Bayesianism as much as explanationism. Let us review these steps in turns.

Consider the famous effect noted by Tversky and Kahneman called the *conjunction fallacy*, which occurs when one judges the probability of a conjunction of two events as being higher than its constituents by themselves. This is fallacious as it violates the basic probabilistic principle: the probability of the conjunction cannot exceed the probabilities of its constituents. Furthermore, Tversky and Kahneman observe that the fallacy can be reliably elicited by a certain form of experimental setting. The paradigmatic case comes to known as the Linda problem. The participants were presented with a description of an individual called Linda:

> Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear

---

[9]Peter Lipton, *Inference to the Best Explanation* (Routledge/Taylor; Francis Group, 2004), 108.

demonstrations.

They were then asked to rate the following statements in terms of probability:

1. Linda is a teacher in elementary school.
2. Linda works in a bookstore and takes Yoga classes.
3. Linda is active in the feminist movement. (A)
4. Linda is a psychiatric social worker.
5. Linda is a member of the League of Women Voters.
6. Linda is a bank teller. (B)
7. Linda is an insurance salesperson.
8. Linda is is active in the feminist movement and a bank teller. $(A \wedge B)$.

The options that really matter are A, B, and $A \wedge B$. The description of Linda is intended to depict someone who is representative of a feminist, but not a bank teller. As mentioned, however, the probability of the conjunction, $A \wedge B$, cannot exceed either A or B. Tversky and Kahneman reports that an overwhelming number of participants rated the conjunction as being more likely, thus suggesting that human beings are highly susceptible to the conjunction fallacy. Lipton suggests that effects such as the conjunction fallacy establish his premise that people are bad at probabilistic reasoning in general.

The second step of the argument is more contentious. Lipton claims that what the Linda problem shows is that people are using Inference to the Best Explanation as a heuristic to estimate the probabilities. The thought is that people are treating the description of Linda as an event to be explained, and they are simply ranking the statements in terms of their explanatory virtues. According to this interpretation, the

18

subjects are ranking Linda being a feminist banker as a more probable than just a banker, because the former provides a better story that explains the description.

A case for compatibilism emerges from these two premises: the empirical evidence shows that people are explanationists, and not Bayesians. Lipton then argues that, due to this empirical fact, Bayesians should make concession to explanationists so that it could supplement Bayesian's descriptive shortcomings: Bayesian conditionalization is too difficult for people to carry out, so they have to use mental shortcuts such as inference to the best explanation to approximate it.

Lipton's argument for compatibilism rests heavily on his account of human psychology, but why should we accept it? He has certainty given *one* story that would explain the findings in behavioral economics, but this is one of many possible hypotheses. To defuse Lipton's argument, all Bayesians have to do is to point out that Lipton's explanationist account is by no means necessary. Consider, for instance, the following alternative hypotheses:

1. *representative heuristic*: people often evaluate conditional probability by using similarity as less computationally demanding substitute. So in this story people are rating Linda being a feminist bank teller as being more probably than a bank teller, because a feminist bank teller is perceived to be more similar to Linda's description than a mere bank teller is.

2. *Bayesian confirmation theory*: it has been pointed out that evidence E can confirm more firmly some hypothesis H_1 relative to H_2, even if $P(H_2|E) > P(H_1|E)$. This happens when E raises the probability of H_1 more so than that of H_2, but not enough to exceed it, that is, $P(H_1|E) - P(H_1) > P(H_2|E) - P(H_2)$.

For an example, consider a standard deck of cards. Suppose a card is drawn randomly and we are interested in knowing whether it's a picture card (C) or if it is a picture card that is also a spades ($C \wedge S$). We receive evidence E: the card is black. E will raise the conditional probability of ($C \wedge S$), from 3/52 to 3/26, even though it is irrelevant to C, since it goes vacuously from 12/52 to 6/26. According to the confirmation-theoretic analysis, description of Linda E is confirmationally more favorable to $A \wedge B$ her being a feminist banker than B being a banker alone. Thus, people are mistaking the probability of the two statements with how well E confirms them.

So there are accounts for the Linda problem that seem to be just as viable as Lipton's proposal. One of them is even within the Bayesian epistemological framework. Now, I am not trying to argue for one of the alternative hypotheses. What I am suggesting is that Lipton's interpretation is by no means necessary, so his point about the data fitting explanationism better is unfounded.

Not only do I think Lipton's argument from human psychology unconvincing, but I also believe that his argument misses the mark *even if* we grant his hypothesis. In other words, even if the IBE hypothesis were true, it still does not show that Bayesians need help from explanationists. Clearly people are not naturally Bayesians, in the strict sense of the term—constantly using conditionalization to precisely update their beliefs in an unbounded manner. Such a feat is too computationally expensive and mathematically intractable for machines. But I do not think that Bayesian epistemology has ever pretended to give a descriptive model of human behavior. It is clearly a normative project, so I fail to see why Bayesians ought to be moved by

Lipton's argument.

My dismissal of Lipton's argument may sound as though I am motivating an anti-naturalistic and anti-scientific position. Let me be clear that my rejecting Lipton's argument from psychology is *not* to say that epistemology should be contained to the airy heights of ideal normative theories. I disagree with Lipton, not because science is irrelevant, but because I disagree with the idea that a theory of epistemology should aim at a descriptively accurate model of human cognition—armchair epistemology makes poor cognitive science. Still, a plausible epistemology ought to provide *practicable* guidance to any intellectually sincere inquirer's quest of knowledge and rationality. As sincere inquirers, we should recognize that we human beings can make systematic errors in thinking. In this context, the failings of the human mind is relevant insofar as revealing the commons pitfall we should avoid, but this *presupposes* a normative theory, which, I think, should be modeled after the successes of our epistemic endeavors, instead of our failures.

2.3   Objective Bayesianism and Explanationism

There is a reading that allows Okasha to sidestep the problem with catchall, but it will not help his claim that PIBE can be can an ally of Bayesianism. This is because this alternative is to say that the example does not involve conditionalization at all. The idea is that if we single out two hypotheses for comparison, we can ascertain which hypothesis would have a higher posterior, without needing the catchall. To see this, consider that

$$\frac{P(H_L|E_S)}{P(H_M|E_S)} = \frac{P(H_L)P(E|H_L)}{P(E_S)} \times \frac{P(E_S)}{P(H_L)P(E_S|H_L)}$$
$$= \frac{P(H_L)P(E_S|H_L)}{P(H_M)P(E_S|H_M)}$$

In other words, to calculate the ratio of the two hypotheses, we do not need the catchall, since it would just get canceled out. Okasha's example might well be a good intuitive example for this comparative inference, but that's besides the point, since he's supposed to be showing us that PIBE can work with conditionalization, which is required for a Strict Bayesianism. Salmon suggests this as a more realistic approach to Bayesian epistemology, but Earman makes it clear that the downgrade Salmon proposes is not acceptable for any Bayesian, since the lack of actual posterior probabilities has rendered most Bayesian methods unworkable. Thus, while this interpretation of Okasha's example eschews the problem of catchall, it does not show that PIBE can work consistently with conditionalization, which makes the example irrelevant.

or essentially the same expression in the form of a ratio:

$$\frac{P(H_L|E_S)}{P(H_M|E_S)} > 1$$

3 Objective Bayesianism

Since we have presented *Subjective* Bayesianism, it seems natural to contrast it with the *objective* form of Bayesianism:

> *Objective Bayesianism*: Priors ought to be determined with objective principles.

In other words, there *are* rational degrees of beliefs, in addition to the subjective ones, and only rational ones are admissible for conditionalization. This usually entails that the priors must be based on some principle, the earliest example of which is Laplace's "Principle of Non-Sufficient Reason":

> The theory of chances consists in reducing all events of the same kind to a certain number of equally possible cases, that is to say, to cases whose existence we are equally uncertain of, and in determining the number of cases favourable to the event whose probability is sought. The ratio of this number to that of all possible cases is the measure of this probability, which is thus only a fraction whose numerator is the number of favourable cases, and whose denominator is the number of all possible cases.[10]

In other words, this principle dictates that when we are in complete ignorance about the out of some event X, the probability *ought* to be equally distributed among all possibilities. For example, if in front of me is a coin I have never seen before in my life, I should believe that the probability of the coin landing on heads ought to be the same as the probability of it landing on tails, which is 1/2. Note that the principle

---

[10]Pierre-Simon Laplace, *Philosophical Essay on Probabilities* (Spring, 1995), 4.

says that in such a situation, the only justifiable prior is 1/2—any other degree would be irrational. The Principle of Non-sufficient Reason is not the only way for a degree of belief to be justified. Another principle is *Calibration*, that says that the rational degree of belief is constrained by empirical evidence.[11]

I shall return to Objective Bayesianism in the discussion of Jonathan Weisberg's attempt to situate explanationism within Objective Bayesianism, but I will point out here that in our context—the Bayeesianism-explanationism compatibility debate—the talk of "Objective Bayesianism" is prone to confusion. The problem is, simply put, that people who seriously claim to be "Objective Bayesian" often do not believe in persistent conditionalization. The reason is that the constraints imposed by the acceptance of both persistent conditionalization and Objective Bayesianism is simply to stringent. For instance, the Principle of Indifference would require us to know all of the possible n outcomes of the experiment, otherwise we cannot assign the probability of $1/n$ to each of them. This means that conditionalization is simply impossible since there is no admissible prior, and if Objective Bayesians are supposed to adhere to Strict Bayesianism, then it would appear that there would be beliefs that cannot be revised rationally, even if we have relevant evidence. Thus, Strict Bayesianism plus Objective Bayesian is a rather extreme position. While it would be incorrect to call J. M. Keynes' view Bayesian, it is arguably closest to what we would call Objective Bayesian here, but he also holds that very few rational priors can actually be known, so many if not most of our problems cannot be given a precise probabilistic analysis. Another example is Jon Williamson, who calls him self an "Objective Bayesian." However, he wholly

---

[11] Jon Williamson, "Objective Bayesianism, Bayesian Conditionalisation and Voluntarism," *Synthese* 178, no. 1 (2011): 67–85 68

rejects conditionalization.[12] In our terminology, he would not be considered as an Objective Bayesian in the strict sense. Going forward, we will need two different terms for these two views. I will do so in our discussion of the compatibility of Objective Bayesianism and explanationism.

4   Conclusion

When explanationists cite van Fraassen's anti-IBE argument, it may sound as though van Fraassen takes Bayesianism as the correct position by fiat, so that any position that contradicts it is *de facto* an inviable one. Indeed, this is how many explanationists read him. Take Lipton for instance:

> In its simplest form, the threatening argument says that Bayesianism is right, so Inference to the Best Explanation must be wrong.[13]

But given our discussion of voluntarism, this assessment of the situation is not quite right, at least in the original context of the argument. The voluntarist argument is not that it is irrational to use conditionalization or IBE, but that (1) they are not rationally compelling in and of themselves, and (2) repacking IBE as some probabilistic rule is inconsistent with Bayesian conditionalization.

Harman, Gilbert. "Knowledge, Inference, and Explanation." *American Philosophical Quarterly* 5, no. 3 (1968): 164–73.

Harman, Gilbert H. "The Inference to the Best Explanation." *Philosophical*

---

[12]Jon Williamson, *In Defence of Objective Bayesianism* (Oxford University Press, 2010), 71.
[13]Lipton, *Inference to the Best Explanation*, 104.

Review 74, no. 1 (1965): 88–95.

Laplace, Pierre-Simon. *Philosophical Essay on Probabilities.* Spring, 1995.

Lipton, Peter. *Inference to the Best Explanation.* Routledge/Taylor; Francis Group, 2004.

Okasha, Samir. "Van Fraassen's Critique of Inference to the Best Explanation." *Studies in History and Philosophy of Science Part A* 31, no. 4 (2000): 691–710.

Williamson, Jon. *In Defence of Objective Bayesianism.* Oxford University Press, 2010.

———. "Objective Bayesianism, Bayesian Conditionalisation and Voluntarism." *Synthese* 178, no. 1 (2011): 67–85.