Kou Okada
April 2023

An Inferentialist Semantics for Bayesian Confirmation Theory

In Bayesian confirmation theory, probability functions are often taken to operate on an algebra of propositions that describe a given theory or evidence (Popper 1934; Sprenger and Hartmann, 2019). So, it is somewhat surprising that there is no substantial literature that connects analyses of sentential meaning (i.e. propositional semantic content) to Bayesian scientific reasoning. In this paper, I begin to articulate a semantic foundation for Bayesianism based on Brandom's (1994) normative inferentialism. In particular, I hope to show that adopting an inferentialist semantics for propositions involved in scientific reasoning allows Bayesians to overcome two key issues: i) the supposed incompatibility between Bayesian confirmation theory and Inference to the Best Explanation (IBE) and ii) the problem of old evidence.

In section 1, I begin by outlining Bayesianism and IBE. I then consider Van Fraassen's (1989) attack on IBE (which entails that Bayesianism and IBE are incompatible) and review a compatibilist response by Okasha (2000). While I find Okasha's response promising, the notion of 'explanatory consideration,' integral to the practice of IBE, is left too vague. In section 2, I lay the foundations for my proposal by reviewing the main tenets of Brandom's (1994) normative inferentialism. Then, in section 3, I flesh out Okasha's proposal by analyzing explanatory power (of theory on evidence) in terms of expressive power (of the inferential network of propositions as a whole). Finally, as an added bonus, in section 4, I show that analyzing explanatory power in terms of expressive power dissolves the problem of old evidence. The upshot is that an inferentialist understanding of propositions involved in scientific reasoning allows for the

rational credentials of common scientific practices—namely deploying IBE and updating the credence of a novel theory using old evidence—to be vindicated.

I. Incompatibility between IBE and Bayesianism?

In this section, I will first review the central features of Bayesian confirmation theory and the schema of Inference to the Best Explanation. Then, I sketch van Fraassen's (1989) influential critique of IBE which entails that Bayesianism and IBE are incompatible. Finally, I will outline and assess a compatibilist response to van Frassen by Okasha (2000). The main takeaway from this section is that although Okasha's compatibilism is on the right track, more needs to be said about the exact notion of explanatory power if we are to fully vindicate the rationality of IBE.

1.

Bayesian confirmation theory aims to describe "scientists as having degrees of belief which are updated by Bayesian Conditionalization," in order to better understand "how [scientists] reason about their theories, why they accept some and reject others" (Sprenger and Hartmann, 2019, p. 24).[1] It has a normative ambition, too: if updating degrees of beliefs, or credences, according to Bayesian Conditionalization is the best (or the only rational) way of doing so, as some Bayesians believe, scientists *ought* to update credences in their theories according to Bayesian Conditionalization. Though approaches to Bayesianism are diverse and varied, three general claims are generally accepted by those that identify as Bayesians:

[1] This is not the only way to understand the main drive of Bayesian confirmation theory. Some philosophers argue that "scientific reasoning is essentially reasoning in accordance with the formal principles of probability" (Urbach & Howson, xvii, 1993), so as to give Bayesianism a central role in science. But I, like Sprenger and Hartmann, find this claim too strong; rationality should not be understood so narrowly, and in practice, there are diverse (and strikingly successful!) ways that scientists reason.

> **Static Dimension**: Rational degrees of belief of an agent are represented by probability functions.
>
> **Dynamic Dimension**: Bayesian Conditionalization (or some generalization thereof) prescribes how a rational agent should revise her degrees of belief.
>
> **Action Dimension**: The posterior probability distribution is the rational basis for assessing evidence, accepting hypotheses and making decisions. (Sprenger and Hartmann, 2019, p. 20)

Let's take them in turn. The static dimension affirms the notion of 'degrees of belief,' which captures the intuition that our confidence in the truth of some beliefs are higher than other beliefs. Further, it states that we are able to model this discrepancy between our confidence in various beliefs (or credences) by assigning them a value between 0—i.e., no confidence in its truth—and 1—i.e., absolute confidence in its truth—through a probability function. As such, the static dimension is largely definitional. Next, the dynamic dimension says that rational agents ought to update their credences (as new information are made available) via conditionalization according to Bayes' Theorem, which is given below:

$$p(H|E) = p(H)\,\frac{p(E|H)}{p(E)}$$

As we are interested in modeling scientific reasoning, what is at stake is how we should revise the credence in hypothesis *H* (or theory) given new information, typically evidence *E*. The basic idea behind Bayesian Conditionalization is quite simple: the rational assignment of credence in a scientific hypothesis *H* after learning evidence *E* is the conditional degree of belief in *H* given *E*. The conditional degree of belief in *H* given *E* is what the Bayes' Theorem describes: in order to calculate it, we need the unconditional degree of belief in *H*, unconditional degree of belief in *E*, and the conditional degree of belief in *E* given *H*.

There are several arguments in favor of endorsing the dynamic dimension of Bayesianism. The perhaps most well-known argument is the Dutch-book argument due to Ramsey (1926). The basic idea is this: instrumentally rational action (e.g. betting) is based on degrees of belief, and the only way to avoid a sure loss (e.g. of money) is by updating our course of action (e.g. accepting and declining bets) according to probabilistic degrees of belief based on conditionalization. This argument provides support for the normative claim that rational agents ought to update their credences by conditionalization only; otherwise, in the long run, we are bound to incur a 'sure loss' and thus can be said to be acting irrationally. Indeed, Lewis and Teller (1973) proved that a Dutch-book can be made against anyone who adopts an explicit rule for updating other than Bayesian conditionalization. (Of course, the Dutch-book argument hinges on the idea that a 'sure loss' is irrational *tout court*, and it is far from clear that a diachronic rule that is liable to the Dutch-book is irrational *tout court.*)

Now, for the action dimension. Since we are interested in the notion of confirmation, we can recast the action dimension in the following way: if the posterior probability (i.e., the conditionalized probability of the hypothesis given evidence) is higher than the prior probability (i.e., the unconditional probability of the hypothesis), then we can say that the evidence confirms the hypothesis and thus rationally increase our credence of the hypothesis. Confirmation comes in degrees, too: the larger the difference is between the posterior and prior probabilities, the more the evidence confirms the hypothesis (and vice-versa).

Now that we have a rough sketch of Bayesian confirmation theory, let us turn our attention to another form of scientific reasoning: Inference to the Best Explanation (IBE). The term IBE was coined by Harman (1965), but there are many precursors, such as the notion of *abduction* due to C. S. Peirce. IBE is a commonly practiced form of reasoning, both in science

and in ordinary life. The basic idea is very simple: after having gleaned the evidence, the reasoner infers the truth of a given hypothesis on the basis that that hypothesis provides the best explanation of the evidence amongst all the other hypotheses available. For example, when I see that the ground is wet outside (evidence), I may infer that it had just rained (hypothesis). I could plausibly infer that someone emptied buckets of water outside, or that an evil demon is tricking me into hallucinating that the ground outside is wet; however, the hypothesis that it had just rained outside simply provides the best explanation of the evidence given what I know. So, using IBE, I was able to infer from the evidence that the ground is wet outside to the hypothesis that it had just rained.

IBE seems to be an attractive scheme for reasoning because it seems to capture, in intuitive terms, how both ordinary people and many scientists infer hypotheses from evidence (with considerable success). However, many philosophers—especially those that endorse probabilistic reasoning as the (only) rational way to carry out non-deductive (i.e. inductive, or abductive, if you'd like) inferences—find IBE too vague and schematic to be considered rational. I will next outline one such influential attack on IBE due to van Fraassen (1989).

2.

Van Fraassen, a noted advocate of Bayesianism, offers a powerful critique of IBE based on three observations:

> 'There are many charges to be laid against the epistemological scheme of Inference to the Best Explanation. One is that it pretends to be something other than it is. Another is that it is supported by bad arguments. A third is that it conflicts with other forms of change of opinion, that we accept as rational." (van Fraassen, 1989, p. 142)

What, according to van Fraaseen, IBE is 'pretending' to be is what he calls the 'traditional ideal of induction,' which has four distinguishing features. Induction, van Fraassen writes, should be "a rule of calculation that extrapolates from particular data to general (or at least ampliative) conclusions. Part of the ideal is (a) that it is a rule, (b) that it is rationally compelling [. . .] (c) that it is objective [. . .] and finally, (d) that it is ampliative" (*ibid.*, p. 132). Van Fraassen accepts that IBE is a rule and that it is ampliative, but rejects that it is rationally compelling.[2] In particular, van Fraassen thinks that IBE "is a rule that selects the best among the historically given hypothesis" (*ibid.*, p. 143) and because this set of hypotheses could be—and historically speaking usually is—a 'bad lot,' i.e., a set of hypotheses that does not contain the truth. Van Fraassen elaborates, "If I do not believe that the truth lies within the set of hypotheses whose explanatory credentials I am examining, then however well the best hypothesis explains the data, I will not believe it. So IBE 'cannot supply the initial context of belief or opinion within which alone it can become applicable."(*Ibid*., p. 149) Indeed, IBE "cannot be what "grounds" rational opinion" because if I don't already think that I am in the right ballpark—i.e., that my initial set of hypotheses contains one that is true—then IBE doesn't even get off of the ground. Van Fraassen puts it this way: "For me to take it that the best of set X will be more likely to be true than not, requires a prior belief that the truth is already more likely to be found in X, than not," (*ibid*., p. 143).

Does this mean that induction and scientific reasoning at large is doomed? No, van Fraassen thinks:

> "If both induction and IBE fail as rational basis for opinion and expectation of the future, traditional epistemology is indeed in serious difficulty. But rather than proclaim the death

[2] While he does not argue that IBE fails to be objective, it seems clear that van Fraassen is suspicious of the very idea of explanatory considerations, let alone its objectivity. So, it seems that van Fraassen would also reject that IBE meets the objectivity criteria.

> of epistemology, and submit either to an irenic relativism or to sceptical despair, I shall try to show in the next chapter that a new epistemology has been quietly growing within the ruins of the old (as well as show that it issues in a still more drastic critique of IBE)" (*ibid*., p. 131).

Of course, the 'new epistemology' he speaks of is Bayesianism. Even if we follow van Fraassen's line of thinking, there still remains a question: Why is Bayesianism and IBE incompatible? After all, IBE has wide currency amongst scientists in practice and it seems to have gotten us decent enough results so far (cf. Putnam's No Miracles Argument), so it would be to the Bayesian's advantage if we could understand the two as compatible ways of reasoning.

One way of reconciling the two approaches is to model IBE under Bayesianism by somehow accommodating the notion of explanatory considerations within the Bayesian technical machinery. Van Fraassen, however, warns us not to listen to those "who preaches a probabilistic version of Inference to the Best Explanation, whatever the details" (*ibid*., p. 169). In particular, van Fraassen considers and dismisses the suggestion that explanatory factors must, rationally speaking, figure into our decision-making process when confronted with novel evidence. He calls this suggestion the 'objection from retrenchment': "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" (*ibid*., p. 160).

To summarize van Fraassen's main argument briefly: if what we mean by explanatory consideration (in IBE) is reducible to posterior probability, then IBE is just Bayesianism in disguise. On the other hand, we could account for explanatory consideration in our probability function by adding "bonus points" onto the conditionalized posterior probability. But, according to van Fraassen, this is irrational: he proves, instantiating Lewis and Teller's general proof that a diachronic Dutch-book can be constructed for any rule of updating credence that is not Bayesian

conditionalization, that a Dutch-book can be constructed for this. Given the—in my view questionable—assumption that being Dutch-book free is the hallmark of rationality, if IBE is a rule for updating credences, then it is not suitably rational. Thus, according to van Fraassen, "when we scrutinize [IBE's] (rational) credentials, we find it seriously wanting." (*ibid.*, p. 131) We can conclude that, on van Fraassen's account, IBE and Bayesianism are fundamentally incompatible.

3.

Responding directly to van Fraassen (1989), Okasha (2000) provides a compatibilist proposal that models IBE on Bayesian terms. He begins by pointing out that IBE is closely related to the hypothetico-deductive (H-D) account of confirmation, under which a given theory is confirmed in case the evidence complies with the observational implications of the theory. This gives us a reason to be suspicious of van Fraassen's claim since Bayesians endorse H-D, albeit in a qualified way: if a theory T entails evidence E, then it follows from the Bayes' theorem that $p(T/E) > p(T)$, and thus that the evidence confirms the theory. IBE differs from this in two respects, namely that the goodness of theories depend on how much they *explain* rather than *entail* the data and allows for a comparison between competing theories. That said, IBE is still very close to the H-D model of confirmation. So, according to Okasha, "if recent authors are right to think that Bayesianism vindicates the basic H-D idea, then van Fraassen's claim that IBE conflicts with Bayesian requirements looks prima facie implausible." (*Ibid*., p. 702)

Naturally, Okasha thinks that van Fraassen has not found the correct way to represent IBE on Bayesian terms. Recall that van Fraassen's 'retrenchment' idea is to model IBE by 'adding bonus points' from explanatory considerations after the posterior probability is calculated. In contrast to this two-stage evolution, Okasha claims that IBE is better understood

not as a separate rule, but a heuristic embedded within the Bayesian machinery. Consider the following example of IBE:

> "A mother takes her five-year-old child to the doctor. The child is obviously in some distress. On the basis of the mother's information, the doctor forms two competing hypotheses: that the child has pulled a muscle, and that he has torn a ligament; call these H1 and H2 respectively. A keen advocate of IBE, the doctor examines the child carefully, and decides that H2 offers the better explanation of the observed symptoms; she therefore tentatively accepts H2—though she does not believe it outright—and rejects H1. 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 if he has pulled a muscle. Therefore the second hypothesis is preferable." (Ibid., p. 702-703)

According to Okasha, we can construct an intuitive Bayesian model of this instantiation of IBE. Given the background information (i.e. the facts presented by the doctor to justify her reasoning), the unconditionalized prior probability of H2 (i.e. torn ligament) is higher than that of H1 (i.e. pulled muscle) and the probability of the evidence conditional on H2 (i.e. symptoms given the child pulled his muscle) is greater than its probability conditional on H1 (i.e. symptoms given the child tore his ligament). Then, it simply follows from the probability calculus that the posterior probability of H2 given evidence is higher than that of H1 given evidence. So, the doctor's reasoning seem to conform perfectly to both IBE and Bayesian principles. In short, "the correct way of representing IBE," Okasha writes, "views the goodness of explanation of a hypothesis vis-à-vis a piece of data as reflected in the prior probability of the hypothesis P(H), and the probability of the data given the hypothesis P(e/H). The better the explanation, the higher is one or both of these probabilities." (*Ibid*., p. 703) Then, after having calculated the posterior probabilities of all the hypotheses, we can simply endorse the hypothesis with the highest

posterior probability. So, IBE represented in such a way is perfectly compatible with Bayesianism.

An added virtue of Okasha's argument is that it captures the phenomenology of scientific reasoning better than van Fraassen's account, and to that extent, more intuitive: instead of the artificial 'two-step' process of reasoning whereby first, posteriors are calculated and then adding bonus points for explanatory considerations, Okasha's account embeds explanatory considerations within the Bayesian machinery, which presumably better captures how scientists reason in practice.

While I think that Okasha's model heads in the right direction, there is a simple question that is left untouched: why should explanatory considerations increase our unconditionalized prior probabilities of the hypothesis and/or that of the evidence given the hypothesis? While it seems intuitive that the background information should influence the way we set priors, especially in the case presented above, the link between explanatory considerations and the setting of our prior probability is not entirely clear. Okasha seems to be claiming that there is a rational way to set priors according to explanatory considerations, and our phenomenology attests to this; however, so long as the link between explanatory consideration and prior probability is left unclear, explanatory considerations play the role of the unexplained explainer on this model. This, I find, frustrating. In section 3, I attempt to articulate the exact relationship between explanatory considerations and priors by analyzing explanatory power, an epistemic concept, in terms of expressive power, a semantic concept. But first, I will lay the foundations by reviewing the semantic framework I appeal to.

II. Inferentialist Semantics and Expressive Rationality

In this section, I outline Brandom's inferentialist approach to semantics with an aim to explicate his notion of expressive rationality. Expressive rationality, I claim, cannot be captured by a notion of rationality oriented by truth-as-correspondence which is a notion often equated with rationality as such. This will pave the way for section 3, where I analyze the explanatory power (of a given theory on some evidence) in terms of the expressive power of the inferential network (of all propositions we are committed to endorsing) as a whole.

1.

To start, the most general demarcation we can make between various philosophical accounts of linguistic meaning (and mental content), in my view, is that any given account falls under either a broadly referentialist program or an inferentialist one. Referentialist program takes linguistic expressions (and concepts) to refer to something in 'objective reality.' Linguistic expressions then are understood as vehicles for *representing* the world.[3] A classic case would be an understanding of propositional content such that content is determined by its truth-conditions. For instance, the meaning of 'The snow is white' is given by conditions (namely, the snow being white) that make the proposition true.[4] In contrast, inferentialist accounts understand the contentfulness of linguistic expressions to be determined by its use, or their inferential roles, namely in drawing inferences to other linguistic expressions. The meaning of 'The snow is white,' under these accounts, is given by totality of the inferences it allows (such as the inference to 'The snow is not red') and it follows from ('There is fresh snow on the ground').

[3] And mind is taken to be the 'mirror of nature,' to quote Rorty's apt phrase. We can keep drawing the analogy between linguistic meaning and mental content, but since for our purposes we are concerned with propositional content, I will focus exclusively on linguistic expression and henceforth drop the comparison to mental content.
[4] Though I've followed Tarski's Convention T here for simplicity, how we cash out the exact notion of 'truth' is inconsequential so long as we understand truth broadly as truth-as-correspondence.

What sets apart Brandom's (1994) inferentialist program from other inferentialist approaches to meaning, such as Inferential/Conceptual Role Semantics (IRS) advocated by Block, Harman, and others, is its normativism regarding content, understood in social pragmatic terms. Whereas IRS takes the meaning constitutive inferential network to be localized in one mind/individual, meaning, on Brandom's account, is determined by socially shared norms that govern our linguistic, and in particular, discursive, practices. It is in virtue of our mutual recognition of our rational nature within a linguistic community—by attributing normative statuses to each others' utterances (and intentional states)—that propositions, and by extension intentional states and concepts, have the meaning they do. Whereas the dominant view is to give a semantics and then extrapolate the pragmatics on that basis (and, to a lesser extent, give an account of mental content and then derive an account of linguistic meaning), Brandom inverts this order of explanation: in short, "semantics must answer to pragmatics" (*ibid.*, p.83) (and mental content is extrapolated from linguistic meaning).

2.

A central feature of Brandom's inferentialism—one which is key for the purposes of this paper—is its commitment to what he calls "expressive rationality," which consists in "making implicit content-conferring inferential commitments explicit as the contents of assertional commitments."[5] (Brandom, 1988, p. 270) In order to feel the force of this idea, we need to understand that, according to Brandom, norms that govern linguistic, and more broadly social, practices come in two varieties: *implicit* and *explicit*. Explicit norms are those norms that can be articulated as rules, principles, or claims (i.e. propositionally contentful items) that hold a given

[5] Given that the title of his most famous book is "Making it Explicit," it is clear the centrality of this idea in Brandom's philosophy.

performance to a normative standard of correctness/incorrectness. For instance, if I say that "Pluto is a planet," this statement can be held to a correctness standard by appealing to some other propositionally contentful items (e.g. the specific definition of planethood). Implicit norms also determine correctness/incorrectness of a given performance, but these are silent norms that remain implicit in practice—in what is done, rather than in what is said. For instance, if someone says "How are you?" to me and I respond, "Pluto is a planet," I have violated some implicit norm that cannot be expressed in propositionally contentful items.[6]

Back to the idea of expressive rationality. Brandom develops this notion from early Frege's logical inferentialist program.[7] Early Frege conceived of the task of the work in developing logic to be "an expressive one; not to prove something, but to say something." He continues: "Right from the start I had in mind the expression of a content. [...] But the content is to be rendered more exactly than is done by verbal language. [...] Speech often only indicates by inessential marks or by imagery what a concept-script should spell out in full." (*Ibid.*, p. 12-13) Frege's logical notation is designed to make explicit conceptual content by specifying the implicit inferential involvements of such a contentful item.

Brandom's contribution to this discussion is to extend this idea of expressive rationality to all kinds of social practices, in particular, linguistic practices of ordinary language. To elucidate the notion of expressive rationality in action, Brandom gives an example of the pejorative term 'Boche'. The meaning of 'Boche' licenses the inferences from "*S* is German" to "*S* is a Boche" and from "*S* is a Boche" to "*S* is stupid." So, the claim "*S* is stupid" follows deductively from "*S* is German." But this is not an inference, he thinks, we ought to permit since we do not rationally endorse the truth of "*S* is stupid." Expressive rationality, in Brandom's view,

[6] I will admit that this is a bad example because Grice's conversation maxims do in fact try to make this implicit norm explicit in the form of propositionally contentful rules. But I hope that the general idea is clear.
[7] And also from Sellars (1950), but I do not have the space to discuss his contribution.

consists of making the implicit inferential relation explicit, then ameliorating the situation by *harmonizing* rational commitments. This must, according to Brandom, "take the form of an investigation of the ongoing elucidative process, of the "Socratic method" of discovering and repairing discordant concepts, which alone gives the notion of harmony any content. It is given content only by the process of harmonizing commitments, from which it is abstracted." (Brandom, 1994, p. 120) So, in the above case, expressive rationality demands that we repair the concept 'Boche,' or, if much of the meaning-constitutive inferential roles do not hold up against the rational scrutiny—which seems to be the case for 'Boche'—we must eliminate the concept altogether.

Though I do not have the space to elaborate further on Brandom's general semantic approach, I hope to have captured some of its main points for the purposes of this paper (namely the notion of "expressive rationality").

3.

There is, however, an independent concern to be addressed before moving onto the next section: Why should we give an inferentialist semantics, rather than, say a referentialist one, for propositions involved in scientific reasoning? Beyond the fact that inferentialism in and of itself is a serious–and in my view correct–theory of linguistic meaning, inferentialism in its normativist variety lends itself particularly well to determine the content of scientific statements (i.e., of a theory or of some evidence) since it embraces the holism that seems inherent in scientific discourse and practice. By holism, I mean the idea that the content (and potentially the truth) of any particular proposition is somehow dependent on all other propositions in the inferential network. For instance, it is not easy, prima facie, to determine the semantic content of 'electron' without referring to related concepts, such as 'proton' or 'atom,' in the same way it

seems plausible to determine the content of ordinary, or non-scientific, concepts like ‘snow’ (by, say, pointing at snow). The insight that science, and associated propositions and concepts, has a radically holistic character has not been lost on many philosophers of science at least since Quine (1951).

But why should we prefer Brandom’s inferentialism over IRS? Brandom’s inferentialism is particularly well-suited for describing scientific concepts because of their inherently social nature: new norms regarding concept usage are constantly being instituted in the scientific community, and it is difficult to account for constant conceptual change in the inferential network of a single individual. A very straightforward example is the concept of ‘planet.’ The norms regarding the use of the concept ‘planet’ changed in 2006 through the consensus of the scientific community; the previously valid and thus meaning-constitutive inference from “This heavenly body is Pluto” to “This heavenly body is a planet” became no longer valid. If the meaning-constitutive inferential network is ‘in the head,’ as IRS suggests, it is difficult, in my view, to make sense of this conceptual change. The deeply social nature of holism in the sciences suggest that Brandom’s brand of inferentialism is better equipped to give a semantics for its concepts and propositions.

III. Explanatory Power and Expressive Power

Now, the stage is set for articulating my main contribution. I will first flesh my proposal out, then I will respond to what I believe to be the main objection.

1.

Let me reiterate the question I posed at Okasha’s compatibilist proposal at the end of section 1: Why should explanatory considerations increase our unconditionalized prior

probabilities of the hypothesis and/or that of the evidence given the hypothesis? What norm(s) of rationality does such a move follow? So long as the notion of explanatory consideration is left vague, we do not have an answer to either of these questions. In light of this, I proposal to analyze the vague notion of explanatory consideration in *semantic* terms; in particular, that explanatory power, defined as the degree to which the hypothesis explains the evidence, should be understood as the degree to which the expressive power of the language (i.e. the inferential network of propositions and concepts) as a whole increases. The rational norm governing the decision to increase our unconditionalized prior probabilities based on explanatory considerations is that of expressive rationality. Thus conceived, together with Okasha's modeling of IBE in Bayesian terms, we can say that the practice of IBE is governed by both expressive rationality and criteria of rationality used by Bayesians (which we may call 'truth-conducive')—albeit in different circumstances, and this avoids the conflict between the two forms of rationality to arise.

I will use Okasha's doctor example to illustrate my point. The doctor appealed to background considerations in forming her unconditionalized credence in two competing hypotheses and her conditionalized credences in evidence given the hypotheses. She was rationally licensed to do so, because she did so in accordance with expressive rationality: she made explicit the implicit inferential relations between propositions whose truth she is rationally committed to. That is, the fact that she is rationally committed to "Pre-adolescent children very rarely pull muscles," "Pre-adolescent children often tear ligaments," "The boy's leg is swollen," " the symptom are exactly what we would expect if the child has torn a ligament," and countless of other meaning-constitutive inferences implies that the doctor is already (implicitly) highly rationally committed the truth of "This child has torn his ligament." All she did by assigning a

higher credence to the prior of this hypothesis compared to the other hypotheses is to make the implicit degrees of commitment to this claim—as determined by the totality of its inferential roles—explicit.

Now, on to the virtues of my proposal: there are three. The first is quite straightforward: by analyzing explanatory power in terms of expressive power, we derive an account of why determining priors based on explanatory consideration should count as rational: it is expressively rational. Thus, I am able to fill in the gaping gap in Okasha's attempt to reconcile IBE and Bayesianism by giving a rationally-acceptable account of modifying priors based on explanatory considerations.

Second, my proposal fills in the gulf between semantics and epistemology, in particular, between two powerful theories with independent support (i.e., Bayesian confirmation theory and social-pragmatic inferentialism). As I stated in the introduction, it is strange that there is not much work dedicated to giving a semantics of scientific propositions, given that Bayesian probability functions operate on propositions. I have tried to ameliorate the situation. And, if the two theories cohere as well as I think they do, according to the notion of explanatory power defended in this paper, that seems to provide sufficient grounds to increase our credences in both theories at the same time.

Finally, and perhaps most controversially, since our explicit commitments to rules, principles, and claims are propositionally contentful, in principle there is a possibility to give exact qualitative values to explanatory power by modeling our degrees of commitment to each scientific proposition in probabilistic terms (via Bayesian networks, for example). Of course, this is very impractical: how are we to weigh our commitment to "The Earth has the mass of x" against, for example, our commitment to "Homo sapiens evolved from chimpanzees" and

countless other scientific claims/rules? But if holism about scientific propositions is correct, nothing short of modeling the whole weighted system of inferential relations will determine an exact value of the degree of commitment to each individual proposition. This consequence may sound unattractive; but, the virtue consists in that *in principle* it is indeed possible to model expressive rationality, just like truth-conducive rationality, in probabilistic terms.

2.

I will briefly rehearse and respond to what I take to be the main objection to my approach: namely, that the notion of expressive rationality can be reduced to truth-conducive rationality since expressive power of a language is the mark of its truth-conduciveness; thus, in the end, truth-conducive rationality is the ultimate source of rationality from which all semantic and epistemic normativity emanates.

This point is well-taken. It seems intuitive enough that scientific rationality should only be concerned with truth (as correspondence with reality). But consider situations where two influential theories contradict one another in some aspects, for example in the case of Einstein's General Theory of Relativity (GTR) and the results obtained from quantum mechanics. To give a reductive summary, quantum field theory assumes absolute time whereas GTR assumes malleable time. If we are to respect the static coherence constraint of truth-conducive rationality (i.e. $p(A) + p(\sim A) = 1$), then the highest degree of credence we can assign equally to each theory given evidence is 0.5, which is indifference. But clearly, we seem to think that the two theories seem to describe physical reality accurately, albeit in two different ways. Indeed, it would make sense to look for a 'missing link' between the theories, i.e. to harmonize our commitments, rather than to abandon one theory altogether. And this is an essentially expressive, rather than truth-conditional, task. Thus conceived, 'truth-conducive rationality' (a term I coined for the

purposes of this paper) may be a misnomer: of course expressive rationality, in the long run, is and ought to be truth-conducive. Updating credences based on expressive rationality and so-called truth-conducive rationality will, over time, converge.

The upshot is that we now have two complimentary notions of rationality. Thus, IBE and Bayesianism work for different types of rationality. One consequence of this pluralism is that Bayesians must drop the idea that an ideally rational agent must *determine* their credences only according to conditionalization, since they ought to comply with expressive rationality in the case of determining the prior probabilities. However, as a rule for updating credences based on novel evidence, Bayesians can still hold onto the claim that an ideally rational agent must *update* their credences only according to conditionalization.

IV. Coda: The Problem of Old Evidence

In concluding my paper, I would like to point out that adopting a pluralist understanding of the normative source of rationality dissolves certain difficulties associated with Bayesian confirmation theory. I have already, implicitly, shown that the so-called Problem of Priors can be accounted for since we have made explicit the expressive character of rational constraint placed on forming prior probabilities.

The Problem of Old Evidence—another notorious problem for Bayesians—too dissolves. To quickly sketch the problem: scientists often take evidence to support a novel theory, even if the evidence was available before the theory was posed. A famous example was when Einstein took the precession of Mercury's perihelion to support his General Theory of Relativity (despite the fact that this data had been well established by the time he came up with GTR). On a Bayesian model, old evidence (i.e. when p(E)=1) cannot confirm a new theory given the standard

Bayesian definition of confirmation, since the posterior probability is always going to be equal to the prior probability.

But, according to the account I have given, we do not need the Bayesian notion of confirmation to rationally license the increase in our credence in a given hypothesis: we can independently assign credence to the hypothesis by appealing to explanatory considerations. This activity is rationally regulated by the distinctly expressive form of rationality. Indeed, one could say that Einstein was 'harmonizing rational commitments' in theorizing GTR which is the expressive rational activity *par excellence*.

V. Works Cited

Brandom, Robert (1988). Inference, expression, and induction. *Philosophical Studies* 54 (2): 257-285.

Brandom, Robert (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press.

Frege, Gottlob (1879). *Begriffsschrift: Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens.* Halle a.d.S.: Louis Nebert.

Harman, Gilbert H. (1965). The inference to the best explanation. *Philosophical Review* 74 (1):88-95.

Lipton, Peter (1991). *Inference to the Best Explanation*. London and New York: Routledge.

Okasha, Samir (2000). Van Fraassen's Critique Of Inference To The Best Explanation. *Studies in History and Philosophy of Science Pa[rt A* 31 (4):691-710.

Popper, Karl R. (1959). *The Logic of Scientific Discovery*. London, England: Routledge.

Quine, Willard V. O. (1951). Two Dogmas of Empiricism. *Philosophical Review* 60 (1):20–43.

Ramsey, Frank (1926). Truth and probability. In Antony Eagle (ed.), *Philosophy of Probability: Contemporary Readings*. Routledge. pp. 52-94.

Sellars, Wilfrid (1950). Language, Rules and Behavior. In Sidney Hook (ed.), *John Dewey: philosopher of science and freedom, a symposium*. New York: The Dial Press. pp. 289–315.

Sprenger, Jan & Hartmann, Stephan (2019). *Bayesian Philosophy of Science: Variations on a Theme by the Reverend Thomas Bayes*. Oxford and New York: Oxford University Press.

Urbach, Peter & Howson, Colin (1993). *Scientific Reasoning: The Bayesian Approach*. Open Court.

Teller, Paul (1973). Conditionalization and observation. *Synthese* 26 (2):218-258.

Van Fraassen, Bas C. (1989). *Laws and Symmetry*. Oxford, England: Oxford University Press.