Learning & Value Change

J. Dmitri Gallow*

Abstract

Accuracy-first accounts of rational learning attempt to vindicate the intuitive idea that, while rationally-formed belief need not be true, it is nevertheless likely to be true. To this end, they attempt to show that the Bayesian's rational learning norms are a consequence of the rational pursuit of accuracy. Existing accounts fall short of this goal, for they presuppose evidential norms which are not and cannot be vindicated in terms of the single-minded pursuit of accuracy. I propose an alternative account, according to which learning experiences rationalize changes in the way you value accuracy, which in turn rationalizes changes in belief. I show that this account is capable of vindicating the Bayesian's rational learning norms in terms of the single-minded pursuit of accuracy, so long as accuracy is rationally valued.

1 Introduction

Daniel the Democrat is relentlessly partisan. If there is a Republican scandal, no matter the details, Daniel is outraged. If there is a Democratic scandal, no matter the details, Daniel is defensive. This isn't because Daniel regards the actions of either Democrats or Republicans as providing evidence about which actions are permissible—he doesn't think that the fact that a Democrat or a Republican lied, embezzled, cheated on their spouse, or what-have-you, makes it any more or less likely that those particular actions are permissible. Nevertheless, if there is a Democratic scandal, he is disposed to believe that the Democrats' actions were permissible. And if there is a Republican scandal, he is disposed to believe that the Republicans' actions were impermissible. And Daniel is disposed to react in this way no matter the details of those scandals. Meanwhile, Melissa the Moderate is disposed to react to political scandals with the same outrage or lack thereof, whether the offenders are Democrats or Republicans. She sometimes believes that Democrats have acted wrongly, and sometimes believes that Republicans have acted wrongly.

Whatever we think about Daniel's conclusions, that shouldn't stop us from saying that Daniel himself is irrational. There is little to admire in his belief-forming practices.

Forthcoming in *Philosophers' Imprint*

✉: jdmitrigallow@pitt.edu

* I am indebted to Boris Babic, Adam Bjorndahl, Michael Caie, Daniel Drucker, Melissa Fusco, Konstantin Genin, Kevin Kelly, Zoë Johnson King, Harvey Lederman, Teddy Seidenfeld, and two anonymous reviewers for helpful conversation on this material.

And whatever we think about Melissa's conclusions, this shouldn't keep us from saying this about her: her belief-formation practices are more rational than Daniel's. Nevertheless, it could turn out—and let us suppose that it does turn out—that the Democrats really haven't done anything impermissible, and all of the Republicans' scandals truly were outrageous. In that case, Daniel's belief-forming practices would have led him to all and only true moral beliefs about American scandals.

That they did, but they are no more rational for it. One can stubbornly stumble into true opinions, but this does not on its own make those opinions rationally held. Likewise, one can be misled into holding false opinions, but their falsehood on its own does nothing to impugn their rationality. So true belief need not be rational, nor rational belief true. On this, most of us are agreed.[1] Nevertheless, there is a long tradition in epistemology of insisting that there is still some intimate connection between rational belief and true belief—in particular, that rational beliefs are *likely* to be true, whereas irrational beliefs are not. While Daniel the Democrat stumbled onto truth, he was more likely to end up with false opinions; and while Melissa the Moderate happened upon falsehood, she was more likely to find herself with true opinions.

Such talk is too Delphic twice over. In the first place, the claims themselves are difficult to interpret—which probability measures underlie these likelihood claims? And why should we care about those probability measures? There are, after all, surely *other* probability measures which give high probability to Daniel's opinions being true. In the second place, the claims are insufficiently motivated—why should we think that Daniel's opinions are not likely to be true? The two concerns are related; if the claim about likelihood is just a claim about objective chances, for instance, then we might think that Democrats are objectively less likely to engage in scandalous behavior than Republicans, in which case, Daniel's belief forming practices are quite likely to lead to truth.

Over the last two or so decades, authors writing under the banner of 'accuracy-first epistemology' have begun to put some meat on the bones of this too-Delphic view of the relationship between rationality and truth.[2] According to accuracy-firsters, what it is to be epistemically rational is, roughly, to pursue accuracy in a manner that would be prudentially rational, were you concerned only with the accuracy of your beliefs. Accuracy-first epistemologists therefore presuppose a form a epistemic consequentialism according to which the sole epistemic good is nearness to truth, or accuracy. Their goal is to *derive* all other epistemic norms from 1) the axiological claim that beliefs are epistemically valuable to the extent that they are accurate; and 2) general consequentialist deontic norms like 'acts are rational if they maximize expected value' and 'it is irrational to take an act that is guaranteed to be worse than another available act'. The

[1] You may disagree; for you may think that belief is rational if and only if it is knowledge. False beliefs cannot be knowledge, and so false beliefs cannot be rational. If you think this, you should still accept that there is *some* important difference between what I call false but rational belief and what I call false and irrational belief. Perhaps you call the former beliefs 'blameless' or 'reasonable' to distinguish them from the latter. If so, then you think that reasonable belief need not be true, nor true belief reasonable. And you face the question of what connection, if any, there is between reasonable belief and truth. Your question and mine are not so different. Perhaps they can receive the same answer.

[2] In addition to the authors discussed below, see Joyce (1998, 2009), Bronfman (2014), Schoenfield (2015, 2016, 2017a, and 2017b), Easwaran (2013, 2016), and Fitelson et al. (ms).

particular epistemic norms I will be focused on here are the Bayesian's rational learning norms (PROBABILISM and CONDITIONALIZATION, to be introduced in §2 below). These norms allow the Bayesian to explain why Daniel is irrational—for Daniel violates the Bayesian's rational learning norms. The accuracy-firsters I will be concerned with here like the Bayesian's explanation of Daniel's irrationality, and they wish to show that the Bayesian's rational learning norms follow from the claims that 1) accuracy is the sole epistemic virtue, inaccuracy the sole epistemic vice; and 2) general consequentialist deontic norms like 'it is rational to adopt beliefs which maximize expected epistemic value'.

According to accuracy-first epistemologists, then, what's wrong with Daniel is that he is either not valuing accuracy properly (in a specific technical sense to be explained below), or else he is adopting opinions which are, *by his own lights*, less likely to be accurate than some other set of opinions he could hold. What makes Melissa's more rational is that she values accuracy properly and pursues the beliefs which are, by her own lights, most likely to be true. (Or at least, we can tell a vindicatory story along these lines about Melissa's beliefs, but not Daniel's.) According to these accuracy-firsters, this is the connection between rational opinion and truth: epistemically rational opinions are those adopted by agents who properly value and rationally seek the truth, the whole truth, and nothing but the truth (§3).

My goal here will be to persuade you that the existing accuracy-first justifications of the Bayesian's rational learning norms presuppose substantive evidential norms which are not themselves explained in terms of the rational pursuit of accuracy; indeed, these evidential norms *cannot* be so explained, given the other commitments of accuracy-first epistemology (§4). For this reason, alternative approaches are needed, and I have one to offer (§5). To preview: on existing accuracy-first approaches, the degree to which you take accuracy at various possibilities into account changes after a learning experience. Once you learn E, you will no longer take accuracy at non-E possibilities into account when deciding which doxastic states are best. However, on standard approaches, this change is the *result* of a rational response to evidence. On existing approaches, first, you rationally respond to learning that E by becoming certain that E. Then, once you've done so, you will end up not taking accuracy at non-E possibilities into account at all, since you are certain that those possibilities are not actual. In contrast, the approach developed here reverses the order of explanation. First, learning that E makes it rational for you to stop valuing accuracy at non-E possibilities, and, for this reason, to stop taking accuracy at those possibilities into account when deciding which doxastic state is best. Changes in your beliefs are then a rational response to these new epistemic values. In a slogan: learning rationalizes epistemic value change.

2 Bayesianism

The Bayesian has a diagnosis of what's gone wrong with Daniel. Their diagnosis is this: either Daniel's opinions are not probabilistically coherent, or Daniel is not a conditionalizer. But being probabilistically coherent and being a conditionalizer are both requirements of rationality. So Daniel is not rational. The remainder of this section will spell out this diagnosis in a bit more detail. Readers already familiar with the Bayesian's rational learning norms should feel free to skip ahead to §3; readers who are additionally familiar with the accuracy-first literature should feel free to skip ahead to §4.

2.1 Probabilism

At any given time, t, you will hold opinions about some propositions. Given a set of doxastically possible worlds, $\mathscr{W}$ (these are possibilities, considered as actual, which cannot be ruled out *a priori*), we can represent the propositions about which you are opinionated at t with sets of possibilities $A \subseteq \mathscr{W}$ (intuitively, the set of possibilities at which A is the case). And we can gather together all of the propositions about which you are opinionated at t into a single set $\mathscr{A}_t$, which we may call your *time t agenda*. For technical reasons, I'll assume throughout that the set of worlds $\mathscr{W}$ is finite.

The particular kind of opinions the Bayesian is concerned with are *degrees of belief*, *degrees of confidence*, or *credences*. We can represent your time t degrees of confidence with a *credence function*, c_t. This is a function from $\mathscr{A}_t$ to $[0,1]$. $c_t(A)$ represents how confident you are that the actual world is one of the doxastically possible worlds in the set A. I'll call a triple $< \mathscr{W}, \mathscr{A}_t, c_t >$—consisting of a set of doxastically possible worlds $\mathscr{W}$, a time t agenda $\mathscr{A}_t$, and a time t credence function c_t—a *credal state*.

PROBABILISM is the claim that, at all times, your degrees of belief ought to be probabilities. That is: for all times t, your opinions should be representable with a credal state $< \mathscr{W}, \mathscr{A}_t, p_t >$ such that $p_t(\mathscr{W}) = 1$ and, for any disjoint $A, B \in \mathscr{A}_t$, $p_t(A \cup B) = p_t(A) + p_t(B)$. (A word on notation: throughout, I'll use 'p' as a name for a credence function when I am assuming that that credence function is a probability, and I'll use 'c' when I am not making this assumption.)

2.2 Conditionalization

Given a credence function c, we may define $c(A \mid E) \overset{\text{def}}{=} c(AE)/c(E)$ to be your credence that A, *given that*, or *conditional on*, E. As I will understand it, CONDITIONALIZATION is the claim that you should take your conditional credences to be your guide for belief-revision. In particular, you should be disposed to, upon receiving the total evidence E, revise your beliefs by adopting a new credence function which is your old credence function conditionalized on E.

A bit of notation: I'll use '$c_{t,E}$' to stand for the credence function that you are disposed to adopt at t, upon possessing the total evidence E. (I will suppress the time-indices when they are irrelevant.) Then, CONDITIONALIZATION is the following thesis.

> Conditionalization
> For all times t, and all propositions A, E such that E could be your total evidence at t,
> $$c_{t,E}(A) \overset{!}{=} c_t(A \mid E)$$

(I place an exclamation mark above the equality to indicate that the claim expressed is normative—CONDITIONALIZATION does not say that you *will* be disposed to conditionalize on your total evidence, merely that you *should* be.)

This, then, is the Bayesian's theory of rational learning: your opinions should be representable with a probability function, and you should be disposed to conditionalize those opinions on your total evidence. In brief, the Bayesian's theory of rational learning is that you should be a probabilistic conditionalizer.

Daniel is not rational, according to the Bayesian, because he is not a probabilistic conditionalizer. To see this, suppose—for *reductio*—that Daniel's opinions are probabilistic and that his is disposed to update by conditionalization. If Daniel learns that a Democrat has ϕ-ed, he is disposed to be very confident that ϕ-ing is not wrong. Thus, if c is Daniel's current credence function, then

1. $c_{\text{a Democrat has } \phi\text{-ed}}(\phi\text{-ing is wrong})$ is low.

And, if he learns that a Republican has ϕ-ed, he is disposed to be very confident that ϕ-ing *is* wrong. So

2. $c_{\text{a Republican has } \phi\text{-ed}}(\phi\text{-ing is wrong})$ is high.

Yet, he thinks that whether ϕ-ing is wrong is independent of whether a Democrat or a Republican has ϕ-ed. So,

3. $c(\phi\text{-ing is wrong} \mid \text{a Democrat has } \phi\text{-ed}) = c(\phi\text{-ing is wrong})$, and

4. $c(\phi\text{-ing is wrong} \mid \text{a Republican has } \phi\text{-ed}) = c(\phi\text{-ing is wrong})$.

So, by (1), (2), and our assumption that Daniel is a conditionalizer,

5. $c(\phi\text{-ing is wrong} \mid \text{a Democrat has } \phi\text{-ed})$ is low, and

6. $c(\phi\text{-ing is wrong} \mid \text{a Republican has } \phi\text{-ed})$ is high.

And, by (3), (4), (5), and (6),

7. $c(\phi\text{-ing is wrong})$ is low, and

8. $c(\phi\text{-ing is wrong})$ is high.

which is not possible if c is a probability function (in fact, it's not possible if c is a *function*). So Daniel is not a probabilistic conditionalizer.

The accuracy-firsters I will be concerned with here like the Bayesian's story about why Daniel is irrational. Their goal is to show that the Bayesian's norms follow from i) the axiological claim that accuracy is the sole epistemic value; ii) a claim about how to *properly* value accurracy; and iii) consequentialist deontic norms like 'credences are rational iff they maximize expected epistemic value'. Before getting to the existing attempts to vindicate CONDITIONALIZATION in terms of the single-minded pursuit of accuracy, let me first say a bit more about (i), (ii), and (iii) above. (Again, readers already familiar with accuracy-first epistemology should feel free to skip ahead to §4.)

A	$p(A)$	$c(A)$
$\varnothing$	0	1
$\{w_1\}$	1/2	1/2
$\{w_2\}$	1/2	1/2
$\mathscr{W}$	1	0

TABLE 1: A probabilistic credence p and a non-probabilistic credence c which agree on the singleton propositions $\{w_1\}$ and $\{w_2\}$.

3 Epistemic Value

I'll use '$\mathcal{V}(c, w)$' to stand for the epistemic value of a credence function c, under the indicative supposition that the world $w \in \mathscr{W}$ is actual. According to accuracy-firsters, $\mathcal{V}(c, w)$ is entirely a function of the *accuracy* of c in world w. For instance, one popular measure of the accuracy of the credence function c in world w_i is the Brier measure,[3]

$$\mathcal{B}(c, w_i) \stackrel{\text{def}}{=} -\sum_{w_j \in \mathscr{W}} (\delta_{ij} - c(w_j))^2$$

Here, '$\delta_{i,j}$' is the Kronecker delta function, which is 1 if $i = j$ and is 0 otherwise. Thus, $\delta_{i,j}$ represents the truth-value of the singleton proposition $\{w_j\}$ at the world w_i. One thing to note about $\mathcal{B}$ is that, according to it, the epistemic value of a credence function is entirely a matter of the credence it places in various propositions—in particular, the credence it places in the singleton propositions—and the truth-value of those propositions. It is in this sense that $\mathcal{B}$ evaluates credences solely in terms of their *accuracy*. You don't need to know, for instance, what your *evidence* is at a world w_i in order to know the epistemic value of a credence function c at that world, according to $\mathcal{B}$.

Another thing to note about $\mathcal{B}$ is that it only pays attention to your credence in the singleton propositions, $\{w_j\}$. Your credences in propositions more coarse-grained than this don't enter into its calculation of epistemic value at all. If we are presupposing that c is a probability function, then this makes sense—for, if c is a probability function (and the number of worlds is finite), then c's credence in every proposition is determined by its credence in the singletons. However, one goal of accuracy-first epistemology is to vindicate the norm of probabilism by appeal to considerations of accuracy, and not merely take it for granted. If this is our goal, then the Brier measure of accuracy will not suit our purposes. Take a simple case in which there are two possible worlds, $\mathscr{W} = \{w_1, w_2\}$, and therefore four propositions: $\varnothing, \{w_1\}, \{w_2\}$, and $\mathscr{W}$ itself. Then, compare the probabilistic p and the non-probabilistic c whose credences in those propositions are as displayed in table 1. p and c will have exactly the same epistemic value at all possible worlds according to $\mathcal{B}$. And if $\mathcal{B}$ does not distinguish between p and c, then we could hardly hope to use the epistemic values of $\mathcal{B}$ to mount a defense of p over c.

The solution is to consider, not just c's credence in the *singleton* propositions, but also its credence in all other propositions $A \in \mathscr{A}$. I will call the straightforward gener-

[3] Throughout, I will abuse notation by writing things like '$c(w_j)$' when I mean '$c(\{w_j\})$'.

alization of $\mathcal{B}$ the *quadratic* measure of accuracy, $\mathcal{Q}$.

$$\mathcal{Q}(c,w) \stackrel{\text{def}}{=} -\sum_{A\in\mathscr{A}} (\chi_A(w) - c(A))^2$$

Here '$\chi_A(w)$' is the characteristic function for the proposition A, which takes the value 1 if A is true at w and takes the value 0 if A is false at w. It therefore represents the truth-value of the proposition A in the world w.

There are other ways of measuring accuracy. Rather than looking at the *square* of the difference between your credence and truth-value, we could look at the *absolute value* of this difference, $\mathcal{A}$; we could look at the *Euclidean distance* between your credence and truth-value, $\mathcal{E}$; or we could look at a *logarithmic* measure of the distance between your credence and truth-value, $\mathcal{L}$.

$$\mathcal{A}(c,w) \stackrel{\text{def}}{=} -\sum_{A\in\mathscr{A}} |\ \chi_A(w) - c(A)\ |$$
$$\mathcal{E}(c,w) \stackrel{\text{def}}{=} -\sqrt{\sum_{A\in\mathscr{A}} (\chi_A(w) - c(A))^2}$$
$$\mathcal{L}(c,w) \stackrel{\text{def}}{=} \sum_{A\in\mathscr{A}} \log[\,|\ (1-\chi_A(w)) - c(A)\ |\,]$$

We may use '$\mathcal{V}_c(c^*)$' to represent how epistemically valuable the credence function c^* is, from the standpoint of the credence function c. Accuracy-firsters like Leitgeb & Pettigrew (2010b) say that, if your credence function is a probability, p, then you should evaluate a credence function c by looking at its *expected epistemic value*, where the expectation is taken with respect to p. This follows from general consequentialist norms which say that the choiceworthiness of an act is given by its expected value.[4] If we understand credence functions as epistemic acts, then this norm tells us that the epistemic choiceworthiness of credence functions is given by their expected epistemic value. So, if p is probabilistic, then

$$\mathcal{V}_p(c) \stackrel{!}{=} \sum_{w\in\mathscr{W}} \mathcal{V}(c,w)\cdot p(w). \tag{3.1}$$

(3.1) says that, if your opinions are probabilistic, then you ought to evaluate credence functions according to their expected epistemic value.

Let's say that an epistemic value function is *proper* iff, when evaluating credences according to that function, every probability function views itself as more valuable than every other credence function.[5]

[4] In the final analysis, it should be either *causal* expected value or *evidential* expected value which guides your evaluation of other credence functions. In cases where there's no act-state dependence, these are both equivalent to the expectation presented in the body. There are interesting issues that crop up when we consider cases of act-state dependence, but I won't be engaging with them here. See Berker (2013), Caie (2013), Greaves (2013), Carr (2017), Konek & Levinstein (forthcoming), and Pettigrew (forthcoming) for more discussion.

[5] This property is often called 'strict propriety' and distinguished from weak propriety. $\mathcal{V}$ is weakly proper iff no probability function p views another credence function c as *more* epistemically valuable than it is. That is, for all p, c: $\mathcal{V}_p(c) \leqslant \mathcal{V}_p(p)$.

Propriety
An epistemic value function $\mathcal{V}$ is *proper* iff, for every probability function p and every credence function $c \neq p$,

$$\mathcal{V}_p(c) < \mathcal{V}_p(p)$$

Propriety looms large in the accuracy-first vindication of Bayesianism. If we assume (3.1)—as I will continue to do for the remainder of the paper—then Predd et al. (2009)'s accuracy-first vindication of probabilism requires only the assumption that the measure of accuracy is proper (along with the condition that the measure is continuous). As we'll see below, Leitgeb & Pettigrew (2010b)'s accuracy-first vindication of conditionalization requires only the assumption that accuracy is valued properly.

Both the quadratic $\mathcal{Q}$ and the logarithmic $\mathcal{L}$ are proper measures of accuracy. However, neither the Brier $\mathcal{B}$,[6] the absolute value $\mathcal{A}$, nor the Euclidean $\mathcal{E}$ are proper measures of accuracy. But each of these are somewhat plausible measures of the distance between a credence function and truth-value. Why shouldn't we value accuracy in the ways prescribed by $\mathcal{B}$, $\mathcal{A}$, or $\mathcal{E}$? There are two arguments of which I am aware.[7] (Note: Pettigrew (2016) has an interesting argument for *the quadratic* accuracy measure in particular. As we'll see below, Levinstein (2012) argues for the logarithmic measure in particular. These are arguments for particular measures of accuracy which happen to be proper; however, they are not directly arguments for the property of propriety itself.)

One argument against improper measures of accuracy, deriving from Oddie (1997), appeals to epistemic conservativism. That argument goes like this.

P1. For any probability function, there is some evidence you could have which would make it epistemically permissible to hold that probability function.

P2. If another credence function has at least as high an expected epistemic value as your own, then it is permissible to adopt that credence function, even without receiving any additional evidence.

P3. It is impermissible to change your credence function without receiving any evidence.

C. So, epistemic value must be proper.

Premise 1 could be justified by an appeal to a radical subjectivism, according to which any probability function is permissible in the absence of evidence. Alternatively, it could be justified by noting that, for any probability function, you could have only the evidence that the the objective chances are given by that probability function.[8] Premise 2 is just a statement of epistemic consequentialism. Premise 3 is the premise of epistemic conservativism.

---

[6] In figure 1, the probabilistic p expects the non-probabilistic c to be exactly as Brier accurate as p itself is (it is impossible for their Brier accuracies to differ).

[7] See Pettigrew (2012a)

[8] For an objection to this justification of premise 1, see Hájek (2008); for a response, see Pettigrew (2016).

Another defense, deriving from Joyce (2009) and Gibbard (2008), justifies propriety by an appeal to *immodesty*—which, in this context, is the thesis that rationality requires you to regard your own credences as more epistemically valuable than any other potential credences. This argument uses premise 1 from the argument from epistemic conservativism, and adds the additional premise of immodesty to conclude that epistemic value must be proper.[9]

P1. For any probability function, there is some evidence you could have which would make it epistemically permissible to hold that probability function.

P4. Rationality requires you to regard your own credences as more epistemically valuable than any other potential credences

C. So, epistemic value must be proper.

(By the way, we'll see in §5.3 that a view like the one sketched here demonstrates that both of these arguments are invalid.)

With this machinery, accuracy-firsters go to work vindicating the central norms of Bayesianism. For instance, Predd et al. (2009) show that, so long as accuracy is properly (and continuously) measured, every non-probabilistic credence function will be *accuracy dominated* by a probabilistic credence function, and no probabilistic credence function is likewise accuracy dominated. That is to say: if you violate the Bayesian's norm of probabilism, then there will be some other, probabilistic, set of credences you could adopt which are guaranteed to be more accurate than your own, no matter which state of the world is actual. Assuming that being dominated in accuracy in this way is irrational, it follows that violating the Bayesian's norm of probabilism is irrational.

Because my focus here is on rational *learning*, and not synchronic rational requirements, I won't dedicate any more attention to the accuracy-first vindication of probabilism.[10] In the next section, I'll consider two attempts to justify diachronic belief-revision norms like conditionalization in terms of the rational pursuit of accuracy and accuracy alone.

Before getting to those vindications, let me comment on some accuracy-first vindications of conditionalization which I will not be discussing. Greaves & Wallace (2006) offer an interestingly different argument than the one I will be discussing below. While Leitgeb & Pettigrew (2010a) are attempting to justify the norm which I have called 'conditionalization', Greaves & Wallace are attempting to justify a subtly different norm which we may call 'plan conditionalization'. While the norm I am calling 'conditionalization' says that it is a rational requirement to be disposed to conditionalize on whatever total evidence you might receive, plan conditionalization says that it is a rational requirement to *plan* to conditionalize on, or to adopt the *strategy* of conditionalizing on, whatever total evidence you might receive. We should carefully distinguish dispositions from plans or strategies. An addict may have plans to refuse cigarettes when they are offered and still be disposed to accept cigarettes when

[9] See Pettigrew (2012b).

[10] The interested reader may consult Joyce (1998, 2009) and Predd et al. (2009); an excellent summary and extension of these arguments is provided in Pettigrew (2016).

they offered. You may be disposed to alert the authorities upon discovering an ominous clown living in the sewers, even if it had never occurred to you to plan for such a contingency. In Kavka (1983)'s toxin puzzle, it may be rational, before midnight, to *plan* to drink the toxin, even if it is not rational to be disposed to drink the toxin when morning arrives. Similarly, even if Greaves & Wallace (2006) are able to show that it is rational to *plan* to conditionalize on whatever total evidence you may receive, this will not on its own demonstrate that you should be disposed to conditionalize on your total evidence once the evidence is in, unless we rely upon a norm saying that you must be disposed to honor all rationally-formed plans, whether or not they continue to maximize expected value when the time comes. Not only is such a norm probably false, but it is difficult to see how to justify such a norm to someone concerned with accuracy and accuracy alone; and who doesn't give a damn about plans except insofar as they contribute to the pursuit of accuracy. For that reason, if the vindication of Greaves & Wallace (2006) is to be used as a vindication of the norm to be disposed to actually revise your degrees of belief by conditionalization when you acquire evidence, as opposed to a vindication of the norm to simply plan to do so, then the objections raised below will apply to their vindication, *mutatis mutandis*.[11] (Parenthetically, the approach of Greaves & Wallace additionally faces other objections, which I detail in Gallow (ms). As I explain there, if you think that you might learn something, but might also learn nothing, then the framework provided by Greaves & Wallace will advise you to plan to not respond to any evidence which you may receive.)[12] There is also a recent approach to vindicating conditionalization in Briggs & Pettigrew (ms) which I won't have the space to discuss here.

4 Conditionalization and Accuracy

In this section, we will consider two separate attempts to show that diachronic norms of belief revision, like conditionalization, follow from the rational pursuit of accuracy—the first in §4.1, and second in §4.2. I will try to persuade you that neither attempt meets with success, because both rely upon substantive evidential norms which are not, and may not be, justified in terms of the single-minded pursuit of accuracy.

4.1 Take One

Leitgeb & Pettigrew (2010b) first attempt to vindicate conditionalization by appealing to a norm stating that, if your credences are given by the probability p, then you should be disposed, upon acquiring the total evidence E, to adopt a new credence function with maximal expected epistemic value within those possibilities not ruled out by E. That is, if your epistemic values are given by $\mathcal{V}$ and your credences are given by p, then the credence p_E you are disposed to adopt upon receiveing total evidence E should be whatever credence function c has maximal expected epistemic value within

[11] It is worth noting that Pettigrew (2016) is far less sanguine about the prospects for defending conditionalization than Leitgeb & Pettigrew (2010b), and endorses instead Greaves & Wallace's vindication of plan conditionalization. See Pettigrew (2016, p. 208).

[12] In this connection, see also Hild (1998), Bronfman (2014), and Schoenfield (2017a).

E possibilities—that is, whichever function c maximizes $\sum_{w \in E} p(w) \cdot \mathcal{V}(c, w)$.[13]

$$p_E \overset{!}{=} \arg\max_c \left\{ \sum_{w \in E} p(w) \cdot \mathcal{V}(c, w) \right\} \tag{4.1}$$

It is important to note that, in (4.1), the function in brackets—the function to be maximized—does not consider the probabilistically-weighted epistemic value of c in *all* worlds in $\mathcal{W}$. Rather, it only looks at probabilistically-weighted epistemic value of c in those worlds compatible with the potential evidence E. Why only the worlds compatible with E? Well, because these are the only possibilities which have not been ruled out by your evidence.

The following proposition shows that, if your epistemic values are proper, then the norm (4.1) entails CONDITIONALIZATION.

Proposition 1 (generalized from Leitgeb & Pettigrew 2010b)**.** If $\mathcal{V}$ is proper, then, for any probability p and any E,

$$\arg\max_c \left\{ \sum_{w \in E} p(w) \cdot \mathcal{V}(c, w) \right\} = p(- \mid E)$$

(See the appendix for a proof.) For the norm (4.1) says that you should be disposed, upon acquiring the total evidence E, to adopt whichever credence function maximizes expected epistemic value within E; and Proposition 1 tells us that, if epistemic value is proper, then the credence function which maximizes expected epistemic value within E is your current credence function conditionalized on E. It follows that, if $\mathcal{V}$ is proper, then you should be disposed to conditionalize on whatever evidence you receive,

$$\text{if } \mathcal{V} \text{ is proper, then } p_E \overset{!}{=} p(- \mid E)$$

Assuming that we are persuaded that we should value accuracy and accuracy alone, and that we should measure accuracy properly, it follows that we should be disposed to conditionalize on our evidence.

This argument is compelling. If successful, it elucidates the connection between responding rationally to your evidence and the rational pursuit of accuracy. If we think that CONDITIONALIZATION is a requirement of rationality, then this argument shows why meeting that requirement means that your beliefs are more likely to be accurate. By the lights of your prior credence function, they are more likely to be accurate (amongst the possibilities consistent with your evidence) than any other credences you could have adopted.

But wait—by what right do we exclude the worlds inconsistent with our evidence from consideration when deciding which posterior credence function to adopt? The general decision-theoretic norm of maximizing expected value tells us to choose a credence function c which maximizes *this* function:

$$\sum_{w \in \mathcal{W}} \mathcal{V}(c, w) \cdot p(w)$$

[13] A word on notation: '$\arg\max_x \{f(x)\}$' denotes the value of x which maximizes the function $f(x)$.

It emphatically does *not* tell us to choose a credence function c which maximizes this *other* function:

$$\sum_{w \in E} \mathcal{V}(c, w) \cdot p(w)$$

Here's a tempting thought—one that tempted me for far too long:[14] we are only taking the expectation over worlds in E because we have learned that the worlds in E are the only live possibilities. This appears to be how LEITGEB & PETTIGREW are thinking about things when they describe the learning event in the following way:

> between t and a later time t', [you obtain] evidence that restricts the set of worlds that are epistemically possible for [you] to the set $E \subset \mathscr{W}$ (p. 249)

On this approach, learning goes in two stages: first, upon receiving evidence E, you eliminate worlds which are incompatible with E from the set of doxastically possible worlds $\mathscr{W}$. Second, you use your prior credence distribution over the remaining worlds to pick a posterior credence distribution, according to the norm (4.1).

Two comments on this approach. Firstly, at stage two, your prior credence distribution over the remaining worlds will not be a probability distribution so long as you were not already certain of E. So $\sum_{w \in E} p(w) \cdot \mathcal{V}(c, w)$ will not, contrary to my earlier loose talk, be an expectation. But the usual reasons for taking functions like this to measure the choice-worthiness of options (the representation theorems, for instance) rely upon p being a probability and the function therefore being an expectation. So, at the very least, something needs to be said about why this new mathematical function should be taken to measure the choice-worthiness of credence functions.[15]

Secondly, due to stage one, this approach does not ultimately justify CONDITIONALIZATION to someone concerned with accuracy and accuracy alone. Why should you stop regarding the worlds outside of E as epistemically possible, and thereby completely discount the accuracy of your credences at worlds outside of E? The natural answer to that question is: "because those worlds are incompatible with your evidence." This answer relies upon a norm like "do not value accuracy at a world if it is incompatible with your evidence". But this is a distinctively *evidential* norm. And we have done nothing to explain why someone who pursues accuracy alone, and cares not at all about evidence *per se* except insofar as it helps them attain their goal of accuracy, will have reason to abide by this evidential norm. So, if our goal was to derive CONDITIONALIZATION from nothing more than the imperative to rationally pursue accuracy, then we have failed.

In fact, things are worse than this. Not only have we not provided an accuracy-based justification of the evidential norm to remove worlds incompatible with your evidence. If accuracy is measured properly, no such justification can be given. For suppose that you have the prior probabilistic credal state $< \mathscr{W}, \mathscr{A}, p >$, and you are evaluating a credal state $< \mathscr{W}_E, \mathscr{A}_E, c_E >$ which has removed worlds incompatible with the evidence E. How should you evaluate the credence function c_E? The accuracy-firster's answer is: you should evaluate it by its expected accuracy—this is just the norm (3.1). And if we evaluate c_E in this way, and if $\mathcal{V}$ is proper, then it is epistemically *worse*

[14] Thanks to Michael Caie for shaking me from my dogmatic slumber on this point.

[15] See CARR (2017).

than your prior credence function. That is, $\mathcal{V}_p(c_E) < \mathcal{V}_p(p)$. This is just what it means for the epistemic value function $\mathcal{V}$ to be proper. So, from the standpoint of your prior credence function, removing worlds incompatible with your evidence is expected to make you *less* accurate. Valuing credence functions with a proper accuracy measure and evaluating credence functions according to their expected accuracy requires you to reason dogmatically. 'Yes, I have evidence that the actual world is in E. But I expect that responding to that evidence by removing $\neg E$ worlds from $\mathscr{W}$ will make my beliefs less accurate. Accuracy is the *only* thing I value. I don't care at all about, *e.g.*, meeting the constraints placed upon me by my evidence, except insofar as doing so conduces to greater accuracy. And I ought to attempt to maximize expected epistemic value. So I ought to ignore this evidence, and maintain my current credences.'

So I don't see how stage 1 of Leitgeb & Pettigrew's two stage approach to rational learning could be justified to someone concerned with accuracy and accuracy alone. But perhaps I was wrong to think of Leitgeb & Pettigrew as relying on a *norm* like "do not treat a world as epistemically possible if it is incompatible with your evidence". Perhaps we should instead understand the account along the following lines: *what it is* to acquire evidence E *just is* for worlds outside of E to be removed from your credal state. Such changes are not to be rationally evaluated, for they are not rationally evaluable. It is a brute psychological fact about you that, upon having a certain experience, some worlds are removed from your credal state. Conditionalization, then, tells you how to be disposed to adopt new credences once this arational psychological change has taken place. While this understanding still must contend with the first concern raised above, it avoids the second. I don't have anything further to say about this idea beyond what has already been said about similar ideas in the work of Richard Jeffrey (1965)—*viz.*, it is patently epistemically irrational to respond to a winter snowfall by eliminating all worlds in which climate change is not a hoax perpetrated by the Chinese. But, if we deny that the elimination of worlds from your credal state is ever rationally evaluable, then we must deny that this particular elimination of worlds from your credal state is irrational. And I'm inclined to regard that as a *reductio* of an epistemological view.

To sum up: Leitgeb & Pettigrew's justification of conditionalization either relies upon a substantive evidential norm which is not, and may not be, justified to someone who is concerned with the pursuit of accuracy and accuracy alone, or else it is committed to the arationality of patently irrational belief changes.

4.2 Take Two

(4.1) is not the only norm that Leitgeb & Pettigrew (2010b) propose for how you should be disposed to respond to acquiring the total evidence E. They additionally suggest that, upon acquiring the total evidence E, you should be disposed to adopt a new credence function c which maximizes expected accuracy *subject to the evidential constraint that E* be given credence 1 and $\neg E$ be given credence 0. Call the set of credence functions which give credence 1 to E and credence 0 to $\neg E$ '$\mathfrak{E}$'. Then, Leitgeb

& Pettigrew endorse the following norm:[16]

(4.2)
$$p_E \overset{!}{=} \arg\max_{c \,\in\, \mathfrak{E}} \left\{ \mathcal{V}_p(c) \right\} \overset{(3.1)}{=} \arg\max_{c \,\in\, \mathfrak{E}} \left\{ \sum_{w \in \mathscr{W}} p(w) \cdot \mathcal{V}(c, w) \right\}$$

That is to say: you should be disposed, upon learning that E, to pick a new credence function from the set $\mathfrak{E}$ of credence functions meeting the evidential constraints. And, amongst the $c \in \mathfrak{E}$, you should be disposed to pick the one which you regard as best. Given (3.1), this means that you should be disposed to pick the one which maximizes expected epistemic value.

Any justification of conditionalization appealing to this norm is susceptible to precisely the same objection we raised in the previous subsection. Why, if we care for accuracy and accuracy alone, should we care about having a credence function within $\mathfrak{E}$? Of course, if we care about meeting the constraints placed upon us by our evidence, we should care about this. Granted. But, as things stand, we have not said anything about why someone who cares about accuracy alone, and cares not at all about evidence *per se*, should want to select a new credence from $\mathfrak{E}$. And again, if $\mathcal{V}$ is proper, there appears to be little we *can* say about this. For, from the standpoint of your prior credence function, p, all the credence functions in $\mathfrak{E}$ are expected to be less accurate than p itself, so long as p is not already in $\mathfrak{E}$. Again, if you care about accuracy and accuracy alone, if you measure it properly, and you abide by the consequentialist deontic norm (3.1), then you will reason dogmatically and refuse to learn from experience.

So, like the norm (4.1), the norm (4.2) may not be justified in terms of the single-minded pursuit of accuracy. In addition, if we measure accuracy with the Brier accuracy measure, $\mathcal{B}$, (as Leitgeb & Pettigrew, 2010a, argue we should) then the norm (4.2) will not vindicate conditionalization. For the (probabilistic) credence function in $\mathfrak{E}$ which maximizes expected $\mathcal{B}$-value from the standpoint of p is not p conditionalized on E. Rather, it is the function $p(- \,||\, E)$, where, for every $A \in \mathscr{A}$,

$$p(A \,||\, E) \overset{\text{def}}{=} p(AE) + \frac{\#A \cap E}{\#E} \cdot [1 - p(E)]$$

(Here '$\#A \cap E$' is the cardinality of the set $A \cap E$, and likewise for '$\#E$'.) To have a name, let's call '$p(A \,||\, E)$' your credence in A *LP conditionalized* on E. When you conditionalize on E, you first set the credence of all worlds outside of E to zero, and then re-normalize in a way that preserves the ratios between your credences. When you LP conditionalize on E, you first set the credence of all worlds outside of E to zero, and then re-normalize in a way that preserves the *differences* between your credences.

[16] The discussion to follow will make it appear that the norms (4.1) and (4.2) are inconsistent. However, Leitgeb & Pettigrew believe that the norms are consistent. This is in part because they think that there is a difference between learning that E and undergoing a learning experience which makes it rational to give credence 1 to E and credence 0 to $\neg E$. The former involves eliminating worlds inconsistent with E from $\mathscr{W}$, while the latter does not. In the body, I will ignore this distinction (in part because it doesn't make a difference for anything I wish to say about the norms and in part because I'm disinclined to think it's a distinction which makes a difference with respect to how you ought to update your credences).

$$\begin{array}{c} \\ G \\ \neg G \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \left[\begin{array}{cc} 3\% & 2\% \\ 57\% & 38\% \end{array}\right] \end{array} \longrightarrow \begin{array}{c} \\ G \\ \neg G \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \left[\begin{array}{cc} 0\% & 32\% \\ 0\% & 68\% \end{array}\right] \end{array}$$

FIGURE 1: The right-hand-side distribution is the result of LP conditionalizing the left-hand-side distribution on the proposition R. (Notice that the *difference* between your credence in $R \cap \neg G$ and your credence in $R \cap G$ has been preserved.)

LEITGEB & PETTIGREW (2010b) offer a proof of the following proposition.[17]

Proposition 2 (LEITGEB & PETTIGREW 2010b)**.** Where $\mathfrak{C}$ is the set of credence functions giving credence 1 to E and credence 0 to $\neg E$ and $\mathfrak{P}$ is the set of probability functions,

$$\arg\max_{c \in \mathfrak{C}, \mathfrak{P}} \left\{ \sum_{w \in \mathcal{W}} p(w) \cdot \mathcal{B}(c, w) \right\} = p(A \,||\, E)$$

If we help ourselves to probabilism and assume that accuracy is to be valued with $\mathcal{B}$, then it follows from (4.2) and **Proposition 2** that we should be disposed, upon acquiring the total evidence E, to LP conditionalize on E.

$$p_E \stackrel{!}{=} p(- \,||\, E) \tag{4.3}$$

(4.3) is not defensible. To co-opt an example from LEVINSTEIN (2012): suppose that I've hidden a prize behind either the left door or the right door, though you don't know which. To begin with, you are 60% confident that the prize is behind the left door and 40% confident that it is behind the right door. Incidentally, you are also 5% confident that ghosts exist. But you don't see any relation between these propositions. You think that whether the prize is behind the left or right door is independent of whether or not ghosts exist. Then, I reveal that the prize is behind the right door. If you then LP conditionalize on the proposition that the prize is behind the right door, you will update your credences as shown in figure 1. (There, 'L' is the proposition that the prize is behind the left door, 'R' the proposition that it is behind the right, and 'G' the proposition that ghosts exist.) So, if you update on this new information with LP conditionalization, your credence that ghosts exist will more than sextuple. You'll end up 32% confident that ghosts exist. There are other problems with LP conditionalization. For instance, suppose that you think there are two ways ghosts could exist, G_1 and G_2, and, conditional on ghosts existing, you divide your credence equally between these two possibilities—but all else is the same as before. Then, you'll end up *even more* confident that ghosts exist—42%. (See figure 2.) Simply by recognizing more ghost possibilities, the information that the prize is behind the right door ends up confirming that ghosts exist to a greater degree. Twisting the knife any further shouldn't be necessary. Updating with LP conditionalization cannot be seriously defended as a rational requirement.

LEVINSTEIN contends that the problem lies, not with the norm (4.2), but rather with the Brier measure of accuracy, $\mathcal{B}$. Instead, LEVINSTEIN suggests using the following

[17] **Proposition 2** is an immediate corollary of Theorem 5 from LEITGEB & PETTIGREW (2010b).

$$\begin{array}{c} \\ G_1 \\ G_2 \\ \neg(G_1 \vee G_2) \end{array} \begin{array}{cc} L & R \\ \left[\begin{array}{cc} 1.5\% & 1\% \\ 1.5\% & 1\% \\ 57\% & 38\% \end{array}\right] \end{array} \longrightarrow \begin{array}{c} \\ G_1 \\ G_2 \\ \neg(G_1 \vee G_2) \end{array} \begin{array}{cc} L & R \\ \left[\begin{array}{cc} 0\% & 21\% \\ 0\% & 21\% \\ 0\% & 58\% \end{array}\right] \end{array}$$

FIGURE 2: The right-hand-side distribution is the result of LP conditionalizing the left-hand-side distribution on the proposition R.

measure of accuracy:

$$\mathcal{L}'(c, w) \stackrel{\text{def}}{=} \ln[c(w)]$$

(Note that this is not what I earlier called 'the logarithmic measure'.) LEVINSTEIN shows that, if you measure accuracy with $\mathcal{L}'$, and you abide by both probabilism and the norm (4.2), then you will be disposed to conditionalize on whatever evidence you receive. That is, LEVINSTEIN shows that

$$\text{(4.4)} \qquad \arg\max_{c \in \mathfrak{E}, \mathfrak{P}} \left\{ \sum_{w \in \mathscr{W}} p(w) \cdot \ln[c(w)] \right\} = p(- \mid E)$$

So, LEVINSTEIN concludes, if you value accuracy in accordance with $\mathcal{L}'$ and you abide by probabilism and the norm (4.2), then you will be a conditionalizer.[18]

But wait—why do we keep assuming probabilism, by requiring that the credence functions from which we are choosing lie in the set $\mathfrak{P}$ of probability functions? To see why, for both LEITGEB & PETTIGREW and LEVINSTEIN, this restriction is crucial, recall the reason I gave in §3 for rejecting accuracy measures which only consider the credence given to singleton propositions. I noted that those measures say nothing at all about your credences in the more coarse-grained propositions, and so do not distinguish between probabilistic and non-probabilistic credence functions (recall table 1). But both $\mathcal{B}$ and $\mathcal{L}'$ are measures of accuracy which only consider the credence given to singleton propositions. So, even if it were the case that $p(- \mid E)$ were *among* the credence functions which maximized expected $\mathcal{L}'$-value, subject to the evidential constraints, it would not *uniquely* maximize expected $\mathcal{L}'$-value, subject to the evidential constraints, since any other credence function which agreed with $p(- \mid E)$ about the credence assigned to the individual worlds would have precisely the same expected $\mathcal{L}'$-value as it.

This observation should not bother us if we already have an accuracy-based justification for ignoring all non-probability functions. However, that justification relied upon the claim that non-probability functions are accuracy-dominated and probability functions are not. This will not be true if accuracy is measured with $\mathcal{L}'$. For *every* probability function is at least weakly $\mathcal{L}'$-dominated by a non-probability function.[19]

[18] I am fudging things here a bit. The expected $\mathcal{L}'$-value of $p(- \mid E)$ will be undefined, since $p(- \mid E)$ gives credence 0 to all worlds in $\neg E$, and $\ln[0]$ is not defined. More carefully, the claim should be that, as x approaches 1, the credence which maximizes expected $\mathcal{L}'$-value, subject to the constraint that $c(E) = x$ and $c(\neg E) = 1 - x$, approaches $p(- \mid E)$.

[19] c *weakly* $\mathcal{L}'$-dominates p iff, for all worlds $w \in \mathscr{W}$, $\mathcal{L}'(p, w) \leqslant \mathcal{L}'(c, w)$ and, for some world $w \in \mathscr{W}$, $\mathcal{L}'(p, w) < \mathcal{L}'(c, w)$. And c *strongly* $\mathcal{L}'$-dominates p iff, for all worlds $w \in \mathscr{W}$, $\mathcal{L}'(p, w) < \mathcal{L}'(c, w)$.

In fact, every probability function is at least weakly $\mathcal{L}'$-dominated by precisely the same non-probability function. This is the function $c^\dagger$, which gives credence 1 to every proposition. Because it is maximally confident in every proposition, I like to call $c^\dagger$ 'the credulous function'. For every world $w \in \mathscr{W}$, $\mathcal{L}'(c^\dagger, w) = \ln[c^\dagger(w)] = \ln[1] = 0$. And 0 is as high as accuracy goes, according to the measure $\mathcal{L}'$. A probability function can only hope to get an $\mathcal{L}'$-value of 0 if it gives all its probability to a single world. But then, while the probability function will do just as well as the credulous function at that world, the credulous function will do better at all other worlds. So the credulous function will weakly $\mathcal{L}'$-dominate probability functions which place all their probability in a single possibility. And it will strongly $\mathcal{L}'$-dominate any probability function which spreads its probability amongst multiple possibilities, since those probabilities will have a negative $\mathcal{L}'$-value at every possible world.[20]

So, returning to LEVINSTEIN's vindication of CONDITIONALIZATION: we should be seriously bothered if $p(- \mid E)$ is merely one among a collection of credence functions which maximize expected $\mathcal{L}'$-value, subject the evidential constraints. For, if we measure accuracy with $\mathcal{L}'$, then we no longer have accuracy-based reasons to rule out non-probability functions. But things are worse than this. It is not even true that $p(- \mid E)$ is *among* the credence functions which maximize expected $\mathcal{L}'$-value, subject to the evidential constraints. The set of credence function which maximize expected $\mathcal{L}'$-value, subject to the constraint that $c(E) = 1$ and $c(\neg E) = 0$, are those functions which give credence 1 to every singleton proposition, and additionally give credence 1 to E and credence 0 to $\neg E$. All such functions meet the evidential constraints and have an expected $\mathcal{L}'$-value of 0, which is, again, as good as $\mathcal{L}'$-value gets.

We could instead measure accuracy with a *proper* logarithmic measure like $\mathcal{L}$, where, recall:

$$\mathcal{L}(c, w) \stackrel{\text{def}}{=} \sum_{A \in \mathscr{A}} \ln[\,|\,(1 - \chi_A(w)) - c(A)\,|\,]$$

$\mathcal{L}$ is both proper and continuous. And, as we saw above, if a measure of accuracy is both proper and continuous, then all non-probability functions will be accuracy-dominated, and no probability function will be accuracy-dominated (by the theorem of PREDD et al. 2009). So we may vindicate probabilism with $\mathcal{L}$. However, $\mathcal{L}$ does not afford a vindication of CONDITIONALIZATION *via* the norm (4.2). To see why, note that, given (3.1),

$$\mathcal{L}_p(c) = \sum_{A \in \mathscr{A}} p(A) \cdot \ln[c(A)] + (1 - p(A)) \cdot \ln[1 - c(A)]$$

If you wish to maximize $\mathcal{L}_p(c)$, then you must choose values $c(A)$ for each $A \in \mathscr{A}$. And, for each $A \in \mathscr{A}$, your choice of $c(A)$ must be the one which maximizes

$$p(A) \cdot \ln[c(A)] + (1 - p(A)) \cdot \ln[1 - c(A)]$$

[20] $\mathcal{L}'$ is often used in statistical applications, where it is taken for granted that the available credence functions are all probabilities. Since it is true that $\mathcal{L}'_p(p^*) < \mathcal{L}'_p(p)$, for any probabilities $p \neq p^*$, the rule is, in these applications, often called 'proper'. This use of 'proper' is, however, importantly different from the definition used by accuracy-firsters, since probabilism is one of the norms they seek to establish, and not one they are willing to take for granted.

For every A, the unique value of $c(A)$ which does this is $p(A)$ (that's just what it is for $\mathcal{L}$ to be proper). Now, if we impose the constraint that $c(E) = 1$ and $c(\neg E) = 0$, this will leave you no leeway with respect to your choice of $c(E)$ and $c(\neg E)$. So those choices will be taken out of your hands. However, the constraint that $c(E) = 1$ and $c(\neg E) = 0$ does not in any way constrain your choice of $c(A)$ for any $A \neq E, \neg E$. So, for all other propositions, $c(A) = p(A)$ will be the only choice which maximizes expected $\mathcal{L}$-value. This means that, if you value accuracy according to $\mathcal{L}$ and you follow the norm (4.2), then you will meet the constraints imposed by your evidence by changing your credence in the propositions E and $\neg E$, but leaving your credence in all other propositions unchanged. Of course, this will not be a probability function, which is to say: the norm (4.2), together with the proper accuracy measure $\mathcal{L}$, is inconsistent with probabilism (and therefore, inconsistent with a norm telling you to avoid accuracy domination). So we had better not accept both (4.2) and adopt $\mathcal{L}$ as our epistemic values.[21]

Even if we require your credences to be a probability function by stipulating that $c(A) = \sum_{w \in A} c(w)$ and measuring accuracy with $\mathcal{L}''$, where

$$\mathcal{L}''(c, w_i) \stackrel{\text{def}}{=} \sum_{w_j \in \mathcal{W}} \ln[\,|\,(1 - \delta_{ij}) - c(w_j)\,|\,]$$

the norm (4.2) will still not vindicate CONDITIONALIZATION. And what it does vindicate is no more defensible than LP conditionalization. Return to LEVINSTEIN's example: you are 60% confident that the prize is behind the left door, and 40% confident it is behind the right. Independently, you are 95% confident that there are no ghosts (that is, you have the credences shown on the left-hand-side of figure 1). If you then learn that the prize is behind the right door and you adopt a credence function with maximal $\mathcal{L}''$-value, subject to the constraint that $\sum_{w \in R} c(w) = 1$ and $\sum_{w \notin R} c(w) = 0$, then you will end up with precisely the same posterior credence function as you would if you LP conditionalized on the proposition R—that is, you end up with the posterior credence distribution shown on the right-hand-side of figure 1.[22,23]

To sum up: LEITGEB & PETTIGREW's attempts to vindicate apparently evidential norms in terms of the rational pursuit of accuracy presuppose substantive evidential norms which are not themselves justified by the rational pursuit of accuracy, and which cannot be justified by the rational pursuit of accuracy so long as accuracy is properly measured. Moreover, their second attempt does not vindicate the norm of CONDITIONALIZATION, but rather the indefensible LP conditionalization. And this consequence of the second attempt may not be mitigated by moving to an alternative measure of accuracy, as LEVINSTEIN (2012) suggests. If we wish to understand rational belief as belief formed in the rational pursuit of truth, we should consider alternative approaches.

[21] For exactly the same reason, we had better not accept (4.2) and adopt $\mathcal{Q}$ as our epistemic values, either.

[22] *Proof sketch*: Let $x = c(R \cap G)$ and $y = c(R \cap \neg G)$. The evidential constraints require your credence in $L \cap G$ and $L \cap \neg G$ to both be 0, so you have no choice about these credences, and you are attempting to choose values of x and y which maximize expected $\mathcal{L}''$-value subject to the constraint that $x + y = 1$. The constraint allows us to substitute $1 - x$ in for y. When we do so (and when we ignore the $\mathcal{L}''$-value of your credence in $L \cap G$ and $L \cap \neg G$, which are fixed by the evidential constraints anyhow), the expected $\mathcal{L}''$-value of a choice of x becomes $0.64 \ln[x] + 1.36 \ln[1 - x]$, which is maximized at $x = 0.32$. (To make this proof less sketch, we should include limits in the manner suggested in footnote 18.)

[23] See Theorem 15.1.4 from PETTIGREW (2016).

5 An Alternative Approach

I have a suggestion. To get you in the mood for the suggestion, allow me to provide, at a very general level, an explanation of why the approach of Leitgeb & Pettigrew ran into the kind of objections I raised above.

What Leitgeb & Pettigrew have provided is a model of rational belief. The model consists of a credal state, an epistemic value function, and a dynamical law which says that credences will move in the direction of highest expected accuracy. However, because the epistemic value function is proper, rational belief will always be in equilibrium (so long as we accept probabilism). If we think that rational belief may *change* as the result of a learning experience, then there must be some exogenously imposed change to some component of this model. There are three options for where this exogenous change could originate: the credal state, the dynamics, or the epistemic value function.

In their first attempted vindication of conditionalization (§4.1), Leitgeb & Pettigrew choose to exogenously alter the credal state by removing worlds from $\mathscr{W}$. If we take this tack, then we face a choice: either we say that the exogenous change to the credal state is rationally evaluable or we say that it is not. If we say that it is not, then we incur counterintuitive consequences like the arationality of becoming certain that Peru will invade China after seeing a bag blowing in the wind. If we say that it is, then we will not have succeeded in our project of explaining the rationality of changes in credal states as the result of the rational pursuit of accuracy; there will be some changes in credal states whose rationality is not explained, and could not be explained, by the single-minded pursuit of accuracy.

Alternatively, there could be an exogenous change in the dynamics. We could say that, while most of the time, rational believers attempt to maximize the accuracy of their beliefs, sometimes, they attempt to meet the constraints imposed by their evidence—even when this will lead to them adopting beliefs which they expect to be less accurate than their current beliefs. This is the route which Leitgeb & Pettigrew take in their attempted vindication of LP conditionalization, and the route which Levinstein takes in his attempted vindication of conditionalization (§4.2). We saw that these attempts resulted in indefensible recommendations. But place that concern to the side. To take this route is to abandon the project of accounting for rational learning in terms of the rational pursuit of accuracy. It is to say that epistemic rationality requires you to follow evidential norms even when following those norms conflicts with the imperative to maximize expected accuracy. Perhaps, at the end of the day, this is what we ought to say. Perhaps, if there is sense to be made of the idea that Melissa's method of belief revision is more likely to lead to accurate beliefs than Daniel's, it is not that Daniel's methods are, by his own lights, expected to lead to less accurate beliefs than he otherwise could have adopted, and Melissa's are, by her own lights, expected to lead to the most accurate beliefs possible.

Perhaps. But before we give up on the accuracy-firster's attempt to flesh out the attractive and intuitive idea that rational responses to evidence are more likely to be accurate than irrational responses to evidence, let's consider the final option: an exogenous change to the epistemic value function.

5.1 Epistemic Value Change

Note that an expected accuracy maximizer will not in general take accuracy at all worlds into account equally. If you are an expected accuracy maximizer and you have the probability function p, then you will evaluate the credence function c by considering the epistemic value of c at each world w, $\mathcal{V}(c, w)$, and *weighting* that value by your credence that w is actual. These weights, $p(w)$, provide some measure of the degree to which you take accuracy at world w into account when you evaluate the credence function c. After a learning experience, then, once you have an updated probability, p', you will take accuracy at some worlds into account more, and take accuracy at some worlds into account less, than you did previously. You will now weight accuracy at world w by $p'(w)$, rather than $p(w)$. So, when you learn that E, you will entirely ignore accuracy at non-E possibilities when evaluating credences. On the standard way of thinking about things, this evaluative change is the *result* of a change in credence. It is only *after* you have rationally responded to a learning experience by becoming certain that E that you stop taking accuracy at non-E possibilities into account. The essence of my suggestion is to reverse the order of explanation. It's not that you should stop taking accuracy at non-E possibilities into account because you should be certain that E. Rather, you should be certain that E because you should stop taking accuracy at non-E possibilities into account in your evaluation of credence functions.

In general, new experiences can rationalize shifts in value. The right kind of experience can rationalize shifts in your aesthetic, moral, and prudential values, *e.g.*. The taste of Vegemite can rationalize valuing or disvaluing foods containing Vegemite, and seeing a Jackson Pollock can rationalize valuing or disvaluing abstract art. Such changes in value can be rational or irrational, depending on the nature of the experience. It would, for instance, be irrational to value abstract art *less* after a resplendent experience with a Jackson Pollock. And these changes in value may render certain changes in behavior rational or irrational. It would be irrational to avoid Vegemite after, and entirely because of, a pleasant experience eating Vegemite on toast.

On the proposal I am putting forward, just as some experiences may rationalize a change in your aesthetic, moral, or prudential values, so too may some experiences—in particular, *learning* experiences—rationalize a change in your epistemic values. An experience which carries the evidence that I have hands rationalizes no longer valuing accuracy at possibilities at which I don't have hands (the 'handless possibilities'). And, importantly, on the current proposal, this change in my epistemic values is not the result of any change in my degrees of belief. I don't stop valuing accuracy at the handless possibilities because I've become certain that I have hands. Rather, I stop valuing accuracy at the handless possibilities because *I've learned* that I don't have hands. On the current proposal, learning experiences rationalize changes in epistemic value; and changes in epistemic value rationalize changes in credence.

One common objection I have encountered is the following: on the current proposal, what *reason* could you have for not caring about accuracy at the non-E possibilities? On existing accuracy-first approaches, you stop taking accuracy at non-E possibilities into account because you are certain that those possibilities aren't actual. But if you're not certain that those possibilities aren't actual—if, instead, you think it's quite *likely* that those possibilities are actual—then what reason could you have to stop valuing accuracy at those possibilities? My response: the reason for not valuing accuracy at the non-E possibilities, in spite of the fact that you think the non-E possibilities are likely, is that you have *learned* that those possibilities are not actual. And *that you've learned E* is a sufficient reason to not value accuracy at non-E possibilities, whatever your prior degrees of belief in E happened to be.

Another common objection is that I have provided no story about why experience rationalizes certain changes in epistemic value and not others. This is true, but a precisely analogous objection applies to existing accuracy-first accounts of rational learning. They have provided no story about what evidence experience provides. Whatever we decide about the relationship between accuracy and belief revision, it is incumbent upon the Bayesian to tell us which evidence propositions are provided by an experience. However, I see no reason why such a story should favor existing accuracy-first accounts of rational learning over my alternative. Take, for instance, the view that your total evidence is just the strongest proposition known.[24] On my proposal, to say this is to endorse the following norm: if you know that E, then it is rational to care only about accuracy in those possibilities in which E is true.

This highlights an important feature of the present proposal: though it claims that the rationality of your degrees of belief is entirely a matter of whether you are rationally pursuing accuracy—though it denies that there are any evidential norms directly governing credence—it is consistent with there being substantive evidential norms governing the rational *evaluation* of credences. That is: while it says that evidence constrains how you may rationally value accuracy, it doesn't say that evidence directly constrains the rationality of your credences themselves. Even if the rationality of your epistemic *values* is in part of function of your evidence, the rationality of your *credences* may remain entirely a function of their expected epistemic value. If rational epistemic value is a function solely of the accuracy of credences, then, even if evidence rationalizes certain ways of valuing accuracy, we may still say that whether your credences are rational is entirely a function of their expected accuracy.

That's the proposal, in broad outline. In §5.2, I'll show that there is a natural way of implementing the proposal on which, if you rationally pursue accuracy and accuracy alone, you will be disposed to conditionalize on whatever evidence you receive. In §5.3, we will see that an account such as this shows that the arguments for propriety we encountered back in §3 are invalid.

5.2 Epistemic Value Change and Conditionalization

Suppose that you have a probabilistic credence function p and that you value accuracy with a proper epistemic value function $\mathcal{V}$. Then, suppose that you undergo a learning experience which makes it rational for you to care *not at all* about how accurate your credences are at the worlds $w \notin E$, but which rationalizes no other change in your epistemic values. Then, it will become rational for you to adopt a new credence function $\mathcal{V}_E$, where

$$\mathcal{V}_E(c, w) = \begin{cases} \mathcal{V}(c, w) & \text{if } w \in E \\ \kappa_w & \text{if } w \notin E \end{cases} \tag{5.1}$$

[24] See Williamson (2000)

and κ_w is any constant. To say that, at worlds $w \notin E$, $\mathcal{V}_E(c, w)$ is a constant is just to say that, at worlds inconsistent with your evidence, you value accurate credences just as much as you value inaccurate credences. Which is just to say that, at those worlds inconsistent with your evidence, you do not value accuracy at all.

The credence function which will maximize expected epistemic value for you, once your epistemic values have shifted to $\mathcal{V}_E$, will be the c which maximizes

$$\begin{aligned} \mathcal{V}_{E,p}(c) &= \sum_{w \in \mathscr{W}} p(w) \cdot \mathcal{V}_E(c, w) \\ &= \sum_{w \in E} p(w) \cdot \mathcal{V}(c, w) \ + \ \sum_{w \notin E} p(w) \cdot \kappa_w \end{aligned}$$

The term $\sum_{w \notin E} p(w) \cdot \kappa_w$ is just a constant, independent of our choice of c. So the choice of c which maximizes $\mathcal{V}_{E,p}(c)$ will be whichever c maximizes

$$\sum_{w \in E} p(w) \cdot \mathcal{V}(c, w)$$

And **Proposition 1** assures is that, since $\mathcal{V}$ was proper, the unique c which maximizes this will be $p(- \mid E)$. That is: once you stop valuing accuracy at non-E possibilities, you will no longer see your prior credence function p as maximizing expected accuracy. Instead, you will see p conditionalized on E as maximizing expected accuracy. Assuming that it is rational to adopt credences which maximize expected accuracy, it is rational to conditionalize on E, once you've stopped valuing accuracy at the non-E possibilities (and your epistemic values have not changed in any other way).

5.3 Propriety and Epistemic Value Change

Note that, if we change the value function in this way, it will no longer be proper. There are probability functions which place positive credence in possibilities incompatible with E—like, *e.g.*, your prior credence function. These probabilities will see some other credence as having higher expected $\mathcal{V}_E$-value than they do. So $\mathcal{V}_E$ is not proper. Is this a problem? If so, it is not because of the arguments which have been advanced for propriety.

Consider first the argument from epistemic conservativism:

- P1. For any probability function, there is some evidence you could have which would make it epistemically permissible to hold that probability function.
- P2. If another credence function has at least as high an expected epistemic value as your own, then it is permissible to adopt that credence function, even without receiving any additional evidence.
- P3. It is impermissible to change your credence function without receiving any evidence.
- C. So, epistemic value must be proper.

I see no reason why any of these premises should be inconsistent with the picture of rational learning I have sketched here. We can accept P1 if we think that any credence function is rational in the absence of evidence—nothing I've said has ruled that out. Moreover, this picture of rational learning entails the epistemic consequentialism of P2. And nothing in the present account is inconsistent with P3. On the present proposal, your epistemic value function will never be proper so long as you have evidence. Nevertheless, so long as your *ur-prior* epistemic value function $\mathcal{V}$—the epistemic value function you held in the absence of any evidence—was proper, every rational probabilistic credence function will, at all times, see itself as uniquely maximizing expected accuracy. So, on the current proposal, the only thing which will prompt a change in credence is the acquisition of evidence. So long as you don't receive any evidence, holding onto your current credences will maximize expected accuracy. Nevertheless, your epistemic value function will not be proper. So we can accept both P1 and P2, as well as the epistemic conservativism of P3, without accepting the conclusion. So the argument is invalid.

The reason the argument is invalid is that it has presupposed that your epistemic value function will remain fixed for all time; that your epistemic values may not *change* as a result of a learning experience. And this is exactly the assumption which we are calling into question here. Precisely the same assumption lies behind the immodesty argument for propriety. Recall, that argument utilizes P1 above, and adds the additional premise:

P4. Rationality requires you to expect your own credences to be more epistemically valuable than any other potential credences.

But nothing we've said here is in any conflict with this premise either. If you update your epistemic values in the way I've proposed, then you *will*, at all times, expect your credences to be more valuable than any others. Again, the argument presupposes that your epistemic value function is fixed for all time. If you deny this assumption, neither of these arguments give you any reason to opt for proper measures of accuracy.

Of course, neither do those arguments give any reason to suspect that your *ur-prior* epistemic values—the epistemic values you adopt prior to receiving any evidence—should be proper. This is true, and it is a problem for the current proposal, since the propriety of the ur-prior epistemic values was crucial to its vindication of CONDITIONALIZATION. Fortunately, there are other arguments for particular proper measures of accuracy which the current proposal does not reveal to be invalid. For instance, PETTIGREW (2016) provides an argument for the quadratic accuracy measure $\mathcal{Q}$ which could easily be co-opted and retrofitted to argue that your ur-prior measure of accuracy must be $\mathcal{Q}$. And since $\mathcal{Q}$ is proper, **Proposition 1** assures us that, if your ur-prior epistemic values are given by $\mathcal{Q}$ and, upon receiving evidence, you update your epistemic values in line with (5.1), then the pursuit of maximum expected epistemic value will compel you to conditionalize on your evidence.

6 In Summation

Existing accuracy-first approaches to rational learning wish to tell the following story about what's wrong with Daniel and what's right with Melissa: what's wrong with Daniel is that Daniel is either failing to value accuracy properly (that is, with a proper measure of accuracy), or else he is not pursuing accuracy rationally (that is, in a prudentially rational manner). And what's right with Melissa is that she is valuing accuracy properly and pursuing accuracy rationally. I've argued that the existing approaches cannot ultimately say these things. Their account of rational learning must presuppose substantive rational norms which do not follow from, and in fact are incompatible with, the imperative to value accuracy properly and pursue accuracy rationally.

On the alternative picture I have sketched, we are able to say the following: what's wrong with Daniel is that he is either failing to value accuracy in a rational way, or else he is not pursuing accuracy rationally. What's right with Melissa is that she is both valuing and pursuing accuracy rationally. What's true in the idea that Daniel's beliefs are not likely to be accurate is that he ought to expect that those beliefs will be less accurate than other beliefs he could have held instead. That is, he ought to respond to his experience by coming to value accuracy in such a way that those beliefs are expected to be less accurate than other ones he could have held instead. What's true in the idea that Melissa's beliefs are likely to be accurate is that she ought to expect them to be more accurate than any other beliefs she could have held instead.

References

Berker, Selim. 2013. "Epistemic Teleology and the Separateness of Propositions." *Philosophical Review*, vol. 122 (3): 337–393. [7]

Brier, Glenn W. 1950. "Verification of Forecasts Expressed in Terms of Probability." *Monthy Weather Review*, vol. 78 (1): 1–3. [6]

Briggs, Rachael & Richard Pettigrew. ms. "An Accuracy-Dominance Argument for Conditionalization." [10]

Bronfman, Aaron. 2014. "Conditionalization and not Knowing that One Knows." *Erkenntnis*, vol. 79 (4): 871–892. [2], [10]

Caie, Michael. 2013. "Rational Probabilistic Incoherence." *Philosophical Review*, vol. 122 (4): 527–575. [7]

Carr, Jennifer. 2017. "Epistemic Utility Theory and the Aim of Belief." *Philosophy and Phenomenological Research*, vol. 95 (3): 511–534. [7], [12]

Easwaran, Kenny. 2013. "Expected Accuracy Supports Conditionalization—and Conglomerability and Reflection." *Philosophy of Science*, vol. 80: 119–142. [2]

—. 2016. "Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probability." *Noûs*, vol. 50 (4): 816–853. [2]

Fitelson, Branden, Kenny Easwaran & David McCarthy. ms. *Coherence*. [2]

Gallow, J. Dmitri. ms. "A Tale of Two Accuracy Maximization Arguments." [10]

Gibbard, Allan. 2008. "Rational Credence and the Value of Truth." In *Oxford Studies in Epistemology*, Tamar Gendler & John Hawthorne, editors, vol. 2, 143–64. Oxford University Press, Oxford. [9]

Greaves, Hilary. 2013. "Epistemic Utility Theory." *Mind*, vol. 122 (488): 915–952. [7]

Greaves, Hilary & David Wallace. 2006. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind*, vol. 115 (495): 607–632. [9], [10]

Hájek, Alan. 2008. "Arguments For—Or Against—Probabilism?" *British Journal for the Philosophy of Science*, vol. 59 (4): 793–819. [8]

Hild, Matthias. 1998. "Auto-Epistemology and Updating." *Philosophical Studies*, vol. 92: 321–361. [10]

Jeffrey, Richard. 1965. *The Logic of Decision*. McGraw-Hill, New York. [13]

Joyce, James M. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science*, vol. 65 (4): 575–603. [2], [9]

—. 2009. "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In *Degrees of Belief*, F. Huber & C. Schmidt-Petri, editors, 263–97. Springer, Dordrecht. [2], [9]

Kavka, Gregory S. 1983. "The Toxin Puzzle." *Analysis*, vol. 43 (1): 33–36. [10]

Konek, Jason & Benjamin Anders Levinstein. forthcoming. "The Foundations of Epistemic Decision Theory." *Mind*. [7]

Leitgeb, Hannes & Richard Pettigrew. 2010a. "An Objective Justification of Bayesianism I: Measuring Inaccuracy." *Philosophy of Science*, vol. 77 (2): 201–235. [9], [14]

—. 2010b. "An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy." *Philosophy of Science*, vol. 77 (2): 236–272. [7], [8], [10], [11], [12], [13], [15], [16], [18], [19], [26]

Levinstein, Benjamin Anders. 2012. "Leitgeb and Pettigrew on Accuracy and Updating." *Philosophy of Science*, vol. 79 (3): 413–424. [8], [15], [16], [17], [18], [19]

Oddie, Graham. 1997. "Conditionalization, Cogency, and Cognitive Value." *British Journal for the Philosophy of Science*, vol. 48: 533–41. [8]

Pettigrew, Richard. 2012a. "Accuracy, Chance, and the Principal Principle." *Philosophical Review*, vol. 121 (2): 241–275. [8]

—. 2012b. "An Improper Introduction to Epistemic Utility Theory." In *EPSA Philosophy of Science: Amsterdam 2009*, Henk W. de Regt, Stephan Hartmann & Samir Okasha, editors, 287–301. Springer Netherlands, Dordrecht. [9]

—. 2016. *Accuracy and the Laws of Credence*. Oxford University Press, Oxford. [8], [9], [10], [18], [23]

—. forthcoming. "Making things right: the true consequences of decision theory in epistemology." In *Epistemic Consequentialism*, K. Ahlstrom-Vij & J. Dunn, editors. Oxford University Press, Oxford. [7]

Predd, Joel, Robert Seiringer, Elliot H. Lieb, Daniel N. Osherson, H. Vincent Poor & Sanjeev R. Kulkarni. 2009. "Probabilistic Coherence and Proper Scoring Rules." *IEEE Transations on Information Theory*, vol. 55 (10): 4786–4792. [8], [9], [17]

Schoenfield, Miriam. 2015. "Bridging Rationality and Accuracy." *Journal of Philosophy*, vol. 112 (12): 633–657. [2]

—. 2016. "An Accuracy Based Approach to Higher Order Evidence." *Philosophy and Phenomenological Research*. [2]

—. 2017a. "Conditionalization does not (in general) Maximize Expected Accuracy." *Mind*, vol. 126 (504): 1155–1187. [2], [10]

—. 2017b. "The Accuracy and Rationality of Imprecise Credences." *Noûs*, vol. 51 (4): 667–685. [2]

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press, Oxford. [21]

A Technicalities

Proposition 1. (generalized from Leitgeb & Pettigrew 2010b) *If $\mathcal{V}$ is proper, then, for any probability p and any E,*

$$p(- \mid E) = \arg\max_c \sum_{w \in E} \mathcal{V}(c, w) \cdot p(w)$$

Proof. $p(- \mid E)$ is a probability function. Since $\mathcal{V}$ is proper, $p(- \mid E)$ maximizes expected $\mathcal{V}$-value, where the expectation is taken relative to itself. So

$$\sum_{w \in \mathcal{W}} \mathcal{V}(c, w) \cdot p(w \mid E) = \sum_{w \in E} \mathcal{V}(c, w) \cdot p(w \mid E)$$

is maximized when $c = p(- \mid E)$. If $p(- \mid E)$ maximizes this function, then it will also maximize the function if we multiply it by the factor $p(E)$. So $p(- \mid E)$ will also maximize

$$\begin{aligned} p(E) \cdot \sum_{w \in E} \mathcal{V}(c, w) \cdot p(w \mid E) &= \sum_{w \in E} \mathcal{V}(c, w) \cdot p(w \mid E) \cdot p(E) \\ &= \sum_{w \in E} \mathcal{V}(c, w) \cdot p(w) \end{aligned}$$

□