What is conditionalization, and why should we do it?[*]

Richard Pettigrew

November 5, 2019

Abstract

Conditionalization is one of the central norms of Bayesian epistemology. But there are a number of competing formulations, and a number of arguments that purport to establish it. In this paper, I explore which formulations of the norm are supported by which arguments. In their standard formulations, each of the arguments I consider here depends on the same assumption, which I call Deterministic Updating. I will investigate whether it is possible to amend these arguments so that they no longer depend on it. As I show, whether this is possible depends on the formulation of the norm under consideration.

One of the central tenets of traditional Bayesian epistemology is Conditionalization. There are various formulations of this norm, but they all agree that it concerns the way your credences should change in response to evidence. I spell out the three formulations that I'll consider below. On the first, Conditionalization concerns how you *actually* update your credences when you receive a piece of evidence; on the second, it concerns how you are *disposed* to update when you receive evidence; and on the third, it concerns how you *plan* to update. In this paper, I am concerned not so much with which formulation of the norm is *correct*—after all, they are not incompatible with each other, and some are independent of each other. Rather, I am concerned with which formulation is *justified* by the existing arguments.

I consider three versions of Conditionalization, and four arguments in their favour. For each combination, I'll ask whether the argument can support the norm when it is formulated in that way. In each case, I note that the standard version of the argument relies on a particular assumption, which I call *Deterministic Updating* and which I formulate precisely below. I'll ask

---

whether the argument really does rely on this assumption, or whether it can be amended to support the norm without that assumption.

This is important because Deterministic Updating says that your updating plan or disposition should specify, for any piece of evidence you might receive, a unique way to update on it. But this seems unmotivated, at least from the Bayesian point of view. After all, subjective Bayesianism is a very permissive theory when it comes to your initial credence function, that is, the one you have at the beginning of your epistemic life before you've gathered any evidence. But, once that initial credence function is chosen from the wide array that Bayesianism deems permissible, the theory is very restrictive about how you should update your credences upon receipt of new evidence. We tolerate this discrepancy because the same sorts of argument seem to give us both the permissiveness of Probabilism and the restrictiveness of Conditionalization. But if it turns out that these arguments only give the latter when we make an unmotivated assumption, this spells trouble for Bayesianism.[1]

I don't claim that the four arguments I consider here exhaust the putative justifications of Conditionalization. Besides these, there are decision-theoretic arguments by Savage (1954, Section 3.5), arguments from symmetry considerations by van Fraassen (1989, Section 13.2) and Grove & Halpern (1998), and arguments from a principle of minimal change due to Diaconis & Zabell (1982, Section 5.1) and Dietrich et al. (2016). I focus on the four described here in the interests of space and because these four arguments are naturally grouped together. We might call them the teleological arguments for Conditionalization, for they seek to establish that norm by pointing to ways that updating in the way it demands optimises different aspects of the goodness of your credences, whether that is their pragmatic utility or their epistemic utility. I leave the question of how these alternative justifications of Conditionalization relate to the assumption of Deterministic Updating to future work.[2]

I start in Section 1 by presenting the various formulations of the norm precisely. Then I introduce the four arguments informally. Then, in Section 2, I introduce some of the formal machinery required to state the arguments. Sections 3-6 contain the central results of the paper. In those sections, I work through each of the four arguments in turn, provide its standard presentation, which assumes Deterministic Updating, and then ask whether we can do without that assumption. As we'll see, for one of the arguments, we cannot do without Deterministic Updating; for the other three, if we drop Deterministic Updating, we face a choice—if we go one way, we can justify the three formulations of Conditionalization without assuming Deterministic Updating; if we go the other way, we cannot. In

---

[1]Thanks to an anonymous referee for suggesting that I mention this motivation.

[2]Thanks to an anonymous referee for urging me to clarify the scope of the present paper.

Section 7, I ask what lessons we can learn from these results.

1 Three formulations and four arguments

Here are the three formulations of Conditionalization. According to the first, Actual Conditionalization, the norm governs your actual updating behaviour.

> **Actual Conditionalization (AC)**
>
> If
>
> (i) c is your credence function at t (I'll often refer to this as your prior);
>
> (ii) the total evidence you receive between t and t' comes in the form of a proposition E learned with certainty;
>
> (iii) $c(E) > 0$;
>
> (iv) c' is your credence function at the later time t' (I'll often refer to this as your posterior);
>
> then it should be the case that $c'(-) = c(-|E) = \frac{c(- \,\&\, E)}{c(E)}$.

According to the second, Plan Conditionalization, the norm governs the updating behaviour you would endorse in all possible evidential situations you might face.

> **Plan Conditionalization (PC)**
>
> If
>
> (i) c is your credence function at t;
>
> (ii) the total evidence you receive between t and t' will come in the form of a proposition learned with certainty, and that proposition will come from the partition $\mathcal{E} = \{E_1, \ldots, E_n\}$;[3]
>
> (iii) R is the plan you endorse for how to update in response to each possible piece of total evidence,
>
> then it should be the case that, if you were to receive evidence E_i and if $c(E_i) > 0$, then R would exhort you to adopt credence function $c_i(-) = c(-|E_i) = \frac{c(- \,\&\, E_i)}{c(E_i)}$.

According to the third formulation, Dispositional Conditionalization, the norm governs the updating behaviour you are disposed to exhibit.

---

[3]A partition is a set of exhaustive and mutually exclusive propositions. That is, the disjunction of the propositions is a tautology, and the conjunction of any two propositions is a contradiction.

Dispositional Conditionalization (DC)

If

(i) c is your credence function at t;

(ii) the total evidence you receive between t and t' will come in the form of a proposition learned with certainty, and that proposition will come from the partition $\mathcal{E} = \{E_1, \ldots, E_n\}$;

(iii) R is the plan you are disposed to follow in response to each possible piece of total evidence,

then it should be the case that, if you were to receive evidence E_i and if $c(E_i) > 0$, then R would exhort you to adopt credence function $c_i(-) = c(-|E_i) = \frac{c(- \, \& \, E_i)}{c(E_i)}$.

Next, let's meet the four arguments. Since it will take some work to formulate them precisely, I will give only an informal gloss here. There will be plenty of time to see them in high-definition in what follows.

Diachronic Dutch Book or Dutch Strategy Argument (DSA) This purports to show that, if you violate Conditionalization, there is a pair of decisions you might face, one before and one after you receive your evidence, such that your prior and posterior credences lead you to choose options when faced with those decisions that are guaranteed to be worse by your own lights than some alternative options (Lewis, 1999).

Expected Pragmatic Utility Argument (EPUA) This purports to show that, if you will face a decision after learning your evidence, then your prior credences will expect your updated posterior credences to do the best job of making that decision if they are obtained by conditionalizing on your priors (Savage, 1954; Good, 1967; Brown, 1976).

Expected Epistemic Utility Argument (EEUA) This purports to show that your prior credences will expect your posterior credences to be best epistemically speaking if they are obtained by conditionalizing on your priors (Greaves & Wallace, 2006).

Epistemic Utility Dominance Argument (EUDA) This purports to show that, if you violate Conditionalization, then there will be alternative priors and posteriors that are guaranteed to be better epistemically speaking, when considered together, than your priors and posteriors (Briggs & Pettigrew, 2018).

2 The framework

In the following sections, I will consider each of the arguments listed above. As we will see, these arguments are concerned directly with updating plans or dispositions, rather than actual updating behaviour. That is, the targets of these arguments—the items that they assess for rationality or irrationality—don't just specify how you in fact update in response to the particular piece of evidence you actually receive. Rather, they assume that your evidence between the earlier and later time will come in the form of a proposition learned with certainty (Certain Evidence); they assume the possible propositions that you might learn with certainty by the later time form a partition (Evidential Partition); and they assume that each of the propositions you might learn with certainty is one about which you had a prior opinion (Evidential Availability); and then they specify, for each of the possible pieces of evidence in your evidential partition, how you might update if you were to receive it.

Some philosophers, like David Lewis (1999), assume that all three of these assumptions—Certain Evidence, Evidential Partition, and Evidential Availability—hold in all learning situations. Others deny one or more. For instance, Richard Jeffrey (1992) denies Certain Evidence and Evidential Availability; Jason Konek (2019) denies Evidential Availability but not Certain Evidence; Bas van Fraassen (1999), Miriam Schoenfield (2017), and Jonathan Weisberg (2007) deny Evidential Partition. But all agree, I think, that there are certain important situations when all three assumptions are true; there are certain situations where there is a set of propositions that forms a partition and about each member of which you have a prior opinion, and the possible evidence you might receive at the later time comes in the form of one of these propositions learned with certainty. Examples might include: when you are about to discover the outcome of a scientific experiment, perhaps by taking a reading from a measuring device with unambiguous outputs; when you've asked an expert a yes/no question; when you step on the digital scales in your bathroom or check your bank balance or count the number of spots on the back of the ladybird that just landed on your hand. So, if you disagree with Lewis, simply restrict your attention to these cases in what follows.

As we will see, we can piggyback on conclusions about plans and dispositions to produce arguments about actual behaviour in certain situations. But in the first instance, I will take the arguments to address plans and dispositions defined on evidential partitions primarily, and actual behaviour only secondarily. Thus, to state these arguments, I need a clear way to represent updating plans or dispositions. I will talk neutrally here of an updating rule. If you think Conditionalization governs your updating dispositions, then you take it to govern the updating rule that matches those dispositions; if you think it governs your updating intentions, then

you take it to govern the updating rule you intend to follow.

I'll introduce a slew of terminology here. You needn't take it all in at the moment, but it's worth keeping it all in one place for ease of reference.

Agenda I will assume that your prior and posterior credence functions are defined on the same set of propositions $\mathcal{F}$, and I'll assume that $\mathcal{F}$ is finite and $\mathcal{F}$ is an algebra. We say that $\mathcal{F}$ is your *agenda*.

Possible worlds Given an agenda $\mathcal{F}$, the set of possible worlds relative to $\mathcal{F}$ is the set of classically consistent assignments of truth values to the propositions in $\mathcal{F}$. I'll abuse notation throughout and write w for (i) a truth value assignment to the propositions in $\mathcal{F}$, (ii) the proposition in $\mathcal{F}$ that is true at that truth value assignment and only at that truth value assignment, and (iii) what we might call the omniscient credence function relative to that truth value assignment, which is the credence function that assigns maximal credence (i.e. 1) to all propositions that are true on it and minimal credence (i.e. 0) to all propositions that are false on it.

Updating rules An *updating rule* has two components:

(i) a set of propositions, $\mathcal{E} = \{E_1, \ldots, E_n\}$

   this contains the propositions that you might learn with certainty at the later time t'; each E_i is in $\mathcal{F}$, so $\mathcal{E} \subseteq \mathcal{F}$; $\mathcal{E}$ forms a partition;

(ii) a set of finite sets of credence functions, $\mathcal{C} = \{C_1, \ldots, C_n\}$

   for each E_i, C_i is the set of possible ways that the rule allows you to respond to evidence E_i; that is, it is the set of possible posteriors that the rule permits when you learn E_i; each c' in C_i in $\mathcal{C}$ is defined on $\mathcal{F}$.[4]

Deterministic updating rule An updating rule $R = (\mathcal{E}, \mathcal{C})$ is *deterministic* if each C_i is a singleton set $\{c_i\}$. That is, for each piece of evidence there is exactly one possible response to it that the rule allows.

Stochastic updating rule A *stochastic updating rule* is an updating rule $R = (\mathcal{C}, \mathcal{E})$ equipped with a probability function P. P records, for each E_i in $\mathcal{E}$ and c' in C_i, how likely it is that I will adopt c' in response to learning E_i. I write this $P(R_{c'}^i | E_i)$, where $R_{c'}^i$ is the proposition that says that you adopt posterior c' in response to evidence E_i.

- I assume $P(R_{c'}^i | E_i) > 0$ for all c' in C_i. If the probability that you will adopt c' in response to E_i is zero, then c' does not count as a response to E_i that the rule allows.

---

[4]For ease of exposition, I'll assume throughout that each C_i contains only finitely many credence functions. Similar results hold if we lift this restriction, but their proofs are more involved and these stronger results aren't needed to make our central point here.

- Note that every deterministic updating rule is a stochastic updating rule for which $P(R_{c'}^i|E_i) = 1$ for each c' in C_i. If $R = (\mathcal{E}, \mathcal{C})$ is deterministic, then, for each E_i, $C_i = \{c_i\}$. So let $P(R_{c_i}^i|E_i) = 1$.

Conditionalizing updating rule An updating rule $R = (\mathcal{E}, \mathcal{C})$ is a *conditionalizing rule* for a prior c if, whenever $c(E_i) > 0$, $C_i = \{c_i\}$ and $c_i(-) = c(-|E_i)$.[5]

Conditionalizing pairs A pair $\langle c, R \rangle$ of a prior and an updating rule is a *conditionalizing pair* if R is a conditionalizing rule for c.

Super-conditionalizing updating rule Suppose $R = (\mathcal{E}, \mathcal{C})$ is an updating rule. Then let $\mathcal{F}^*$ be the smallest algebra that contains all of $\mathcal{F}$ and also $R_{c'}^i$ for each E_i in $\mathcal{E}$ and c' in C_i. (As above $R_{c'}^i$ is the proposition that says that you adopt posterior c' in response to evidence E_i.) Then

(a) *R is a weak super-conditionalizing rule for c* if there is an extension c^* of c such that, for all E_i in $\mathcal{E}$ and c' in C_i, if $c^*(R_{c'}^i) > 0$, then $c'(-) = c^*(-|R_{c'}^i)$.

   That is, each posterior to which you assign positive prior credence is the result of conditionalizing the extended prior c^* on the evidence to which it is a response and the fact that it was your response to this evidence.

(b) *R is a strong super-conditionalizing rule for c* if there is an extension c^* of c such that, for all E_i in $\mathcal{E}$ and c' in C_i, $c^*(R_{c'}^i) > 0$ and $c'(-) = c^*(-|R_{c'}^i)$.

   That is, you assign positive prior credence to each posterior and each posterior is the result of conditionalizing the extended prior c^* on the evidence to which it is a response and the fact that it was your response to this evidence.

Super-conditionalizing pair A pair $\langle c, R \rangle$ of a prior and an updating rule is a *weak (strong) super-conditionalizing pair* if R is a weak (strong) super-conditionalizing rule for c.

Let's illustrate these definitions using an example. Condi is a meteorologist. There is a hurricane in the Gulf of Mexico. She knows that it will make landfall soon in one of the following four towns: Pensacola, FL, Panama City, FL, Mobile, AL, Biloxi, MS. She calls a friend and asks whether it has hit yet. It has. Then she asks whether it has hit in Florida. At this point, the evidence she will receive when her friend answers is either F—which says that it made landfall in Florida, that is, in Pensacola

---

[5]Note that a conditionalizing rule for a prior need not be a deterministic updating rule. It need only be deterministic for those possible pieces of evidence to which the prior assigns positive credence.

or Panama City—or $\overline{F}$—which says it hit elsewhere, that is, in Mobile or Biloxi. Her prior is c:

	Panama City	Pensacola	Mobile	Biloxi
c	60%	20%	15%	5%

Her evidential partition is

$$\mathcal{E} = \{F = \textit{Pensacola} \vee \textit{Panama City}, \overline{F} = \textit{Mobile} \vee \textit{Biloxi}\}$$

And here are some posteriors she might adopt:

	Panama City	Pensacola	Mobile	Biloxi
c'_F	75%	25%	0%	0%
$c'_{\overline{F}}$	0%	0%	75%	25%
c°_F	70%	30%	0%	0%
$c^{\circ}_{\overline{F}}$	0%	0%	70%	30%
$c^{\dagger}_F$	80%	20%	0%	0%
$c^{\dagger}_{\overline{F}}$	0%	0%	80%	20%

And here are four possible rules she might adopt, along with their properties:

	F	$\overline{F}$	Det.	Cond.	W./S. Super-cond.
R_1	$\{c'_F\}$	$\{c'_{\overline{F}}\}$	✓	✓	✓
R_2	$\{c^{\circ}_F\}$	$\{c^{\circ}_{\overline{F}}\}$	✓	✗	✗
R_3	$\{c^{\circ}_F, c^{\dagger}_F\}$	$\{c^{\circ}_{\overline{F}}, c^{\dagger}_{\overline{F}}\}$	✗	✗	✓
R_4	$\{c^{\circ}_F\}$	$\{c^{\circ}_{\overline{F}}, c^{\dagger}_{\overline{F}}\}$	✗	✗	✗

We'll see in detail below why R_3 is a strong super-conditionalizing rule for c, but roughly speaking the reason is that it has two properties that are jointly sufficient for being such a rule, as Lemma 2 shows: first, each posterior that R_3 permits assigns maximum credence to the evidence to which it is a response; second, c is a weighted average of those permitted posteriors in which the weights are all positive.

As we will see below, for each of our four arguments for Conditionalization—DSA, EPUA, EEUA, and EUDA—the standard formulation of the argument assumes a norm that I call Deterministic Updating:

Deterministic Updating (DU) Your updating rule should be deterministic.

In what follows, I will present each argument in its standard formulation, which assumes Deterministic Updating. Then I will explore what happens when we remove that assumption.

3 The Dutch Strategy Argument (DSA)

The DSA and EPUA both evaluate updating rules by considering their pragmatic consequences. That is, they look to the choices that your priors and/or your possible posteriors lead you to make and they conclude that they are optimal only if your updating rule is a conditionalizing rule for your prior.

3.1 DSA with Deterministic Updating

Let's look at the DSA first. In what follows, I'll take a decision problem to be a set of options that are available to an agent: e.g. accept a particular bet or refuse it; buy a particular lottery ticket or don't; take an umbrella when you go outside, take a raincoat, or take neither; and so on. The idea behind the DSA is this. One of the roles of credences is to help us make choices when faced with decision problems. They play that role badly if they lead us to make one series of choices when another series is guaranteed to serve our ends better. The DSA turns on the claim that, unless we update in line with Conditionalization, our credences will lead us to make such a series of choices when faced with a particular series of decision problems.

Here, I restrict attention to a particular class of decision problems you might face. They are the decision problems in which, for each available option, its outcome at a given possible world obtains for you a certain amount of a particular quantity, such as money or chocolate or pure pleasure, and your utility is linear in that quantity—that is, obtaining some amount of that quantity increases your utility by the same amount regardless of how much of the quantity you already have. The quantity is typically taken to be money, and I'll continue to talk like that in what follows. But it's really a placeholder for some quantity with this property. I restrict attention to such decision problems because, in the argument, I need to combine the outcome of one decision, made at the earlier time, with the outcome of another decision, made at the later time. So I need to ensure that the utility of a combination of outcomes is the sum of the utilities of the individual outcomes.

Now, suppose c is our prior and $R = (\mathcal{E} = \{E_1, \ldots, E_n\}, \mathcal{C} = \{C_1, \ldots, C_n\})$ is our updating rule. As I do throughout, I assume that c is a probability function, and so is each c' in C_i in $\mathcal{C}$. And I will assume further that, when your credences are probabilistic, and you face a decision problem, then you should choose from the available options one that maximises expected utility relative to your credences.

With this in hand, let's define two closely related features of a pair $\langle c, R \rangle$ that are undesirable from a pragmatic point of view, and might be thought to render that pair irrational. First:

Strong Dutch Strategies $\langle c, R \rangle$ is vulnerable to a *strong Dutch strategy* if there are two decision problems, $\mathbf{d}$ and $\mathbf{d}'$, such that

   (i) c requires you to choose option A from the possible options available in $\mathbf{d}$;

  (ii) for each E_i in $\mathcal{E}$ and c' in C_i, c' requires you to choose X from $\mathbf{d}'$;

 (iii) there are alternative options, B in $\mathbf{d}$ and Y in $\mathbf{d}'$, such that, at every possible world, you'll receive more utility from choosing B and Y than you receive from choosing A and X. In the language of decision theory, $B + Y$ strongly dominates $A + X$.

Weak Dutch Strategies $\langle c, R \rangle$ is vulnerable to a *weak Dutch strategy* if there is a decision problem $\mathbf{d}$ and, for each E_i in $\mathcal{E}$ and c' in C_i, a further decision problem $\mathbf{d}_{c'}^i$ such that

   (i) c requires you to choose A from $\mathbf{d}$;

  (ii) for each E_i in $\mathcal{E}$ and c' in C_i, c' requires you to choose $X_{c'}^i$ from $\mathbf{d}_{c'}^i$;

 (iii) there is an alternative option, B in $\mathbf{d}$, and, for each E_i in $\mathcal{E}$ and c' in C_i, there is an alternative option, $Y_{c'}^i$ in $\mathbf{d}_{c'}^i$, such that (a) for each E_i, each world in E_i, and each c' in C_i, you'll receive at least as much utility at that world from choosing B and $Y_{c'}^i$ as you'll receive from choosing A and $X_{c'}^i$, and (b) for some E_i, some world in E_i, and some c' in C_i, you'll receive strictly more utility at that world from B and $Y_{c'}^i$ than you'll receive from A and $X_{c'}^i$.

Then the Dutch Strategy Argument is based on the following mathematical fact (de Finetti, 1974):

Theorem 1 *Suppose R is a deterministic updating rule. Then:*

  (i) *If R is not a conditionalizing pair for c, then $\langle c, R \rangle$ is vulnerable to a strong Dutch strategy;*

 (ii) *If R is a conditionalizing rule for c, then $\langle c, R \rangle$ is not vulnerable even to a weak Dutch strategy.*

That is, if your updating rule is not a conditionalizing rule for your prior, then your credences will lead you to choose a strongly dominated pair of options when faced with a particular pair of decision problems; if you satisfy it, that can't happen.

Now that we have seen how the argument works, let's see whether it supports the three versions of Conditionalization that we met above: Actual (AC), Plan (PC), and Dispositional (DC) Conditionalization. Since they speak directly of rules, let's begin with PC and DC.

The DSA shows that, if you endorse a deterministic rule that isn't a conditionalizing rule for your prior, then there is pair of decision problems, one that you'll face at the earlier time and the other at the later time, where your credences at the earlier time and your planned credences at the later time will require you to choose a dominated pair of options. And it seems reasonable to say that it is irrational to endorse a plan when you will be rendered vulnerable to a Dutch Strategy if you follow through on it. So, for those who endorse deterministic rules, DSA plausibly supports Plan Conditionalization.

The same is true of Dispositional Conditionalization. Just as it is irrational to *plan* to update in a way that would render you vulnerable to a Dutch Strategy if you were to stick to the plan, it is surely irrational to be *disposed* to update in a way that will render you vulnerable in this way. So, for those whose updating dispositions are deterministic, DSA plausibly supports Dispositional Conditionalization.

Finally, AC. There are various different ways to move from either PC or DC to AC, but each one of them requires some extra assumptions. For instance:

(1) I might assume: (i) between an earlier and a later time, there is always a partition such that you know that the strongest evidence you will receive between those times is a proposition from that partition learned with certainty; (ii) if you know you'll receive evidence from some partition, you are rationally required to plan how you will update on each possible piece of evidence before you receive it; and (iii) if you plan how to respond to evidence before you receive it, you are rationally required to follow through on that plan once you have received it. Together with PC + DU, these give AC.

This is the most common route to AC, and has therefore received the most attention. Miriam Schoenfield (2017), Jonathan Weisberg (2007), Bas van van Fraassen (1999), and Aaron Bronfman (2014) deny (i); Bas van Fraassen (1989) denies (ii); and Richard Pettigrew (2016) denies (iii).

(2) I might assume: (i) you have updating dispositions. So, if you actually update other than by Conditionalization, then it must be a manifestation of a disposition other than conditionalizing. Together with DC + DU, this gives AC.

(3) I might assume: (i) that you are rationally required to update in any way that can be represented as the result of updating on a plan that

you were rationally permitted to endorse or as the result of disposi-
tions that you were rationally permitted to have, even if you did not
in fact endorse any plan prior to receiving the evidence nor have any
updating dispositions. Again, together with PC + DU or DC + DU,
this gives AC.

Notice that, in each case, it was essential to invoke Deterministic Updating
(DU). As we will see below, this causes problems for AC.

3.2 DSA without Deterministic Updating

We have now seen how the DSA proceeds if we assume Deterministic Up-
dating. But what if we don't? Consider, for instance, rule R_3 from our
list of examples at the end of Section 2 above, where I described Condi's
credences concerning the landfall of a hurricane:

$$R_3 = (\mathcal{E} = \{F, \overline{F}\}, \mathcal{C} = \{C_F = \{c_F^\circ, c_F^\dagger\}, C_{\overline{F}} = \{c_{\overline{F}}^\circ, c_{\overline{F}}^\dagger\}\})$$

That is, if Condi learns F, rule R_3 allows her to update from her prior c to
posterior c_F° or posterior $c_F^\dagger$. And if she receives $\overline{F}$, it allows her to update to
$c_{\overline{F}}^\circ$ or to $c_{\overline{F}}^\dagger$. Notice that $\langle c, R_3 \rangle$ violates Conditionalization thoroughly: it is
not deterministic; and, moreover, as well as not mandating the posteriors
that Conditionalization demands, it does not even permit them. The poste-
rior $c(-|F)$ does not appear in C_F and $c(-|\overline{F})$ does not appear in $C_{\overline{F}}$. Can
we adapt the DSA to show that R_3 is irrational for someone with prior c?
As we'll see, the answer is no. The reason is that $\langle c, R_3 \rangle$ is not vulnerable to
a Dutch Strategy.

To see this, I first note that, while R_3 is not deterministic and not a con-
ditionalizing rule for c, it is a super-conditionalizing rule for c. And to
see that, it helps to state the following representation theorem for super-
conditionalizing rules, which we mentioned informally above:

Lemma 2

(i) *R is a weak super-conditionalizing rule for c iff there is, for each E_i in $\mathcal{E}$ and
c' in C_i, $0 \leq \lambda_{c'}^i \leq 1$ with $\sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda_{c'}^i = 1$ such that*

   (a) *for all E_i in $\mathcal{E}$ and c' in C_i, if $\lambda_{c'}^i > 0$, then $c'(E_i) = 1$, and*
   (b) $c(-) = \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda_{c'}^i c'(-).$

(ii) *R is a strong super-conditionalizing rule for c iff there is, for each E_i in $\mathcal{E}$
and c' in C_i, $0 < \lambda_{c'}^i < 1$ with $\sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda_{c'}^i = 1$ such that*

   (a) *for all E_i in $\mathcal{E}$ and c' in C_i, $c'(E_i) = 1$; and*
   (b) $c(-) = \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda_{c'}^i c'(-).$

Now note:

(a) $c_F^\circ(F) = 1 = c_F^\dagger(F)$ and $c_{\overline{F}}^\circ(\overline{F}) = 1 = c_{\overline{F}}^\dagger(\overline{F})$

(b) $c(-) = 0.4c_F^\circ(-) + 0.4c_F^\dagger(-) + 0.1c_{\overline{F}}^\circ(-) + 0.1c_{\overline{F}}^\dagger(-)$

So R_3 is a strong super-conditionalizing rule for c. What's more:

Theorem 3

  (i) *If R is not a weak or strong super-conditionalizing rule for c, then $\langle c, R \rangle$ is vulnerable at least to a weak Dutch Strategy, and possibly also a strong Dutch Strategy.*

  (ii) *If R is a strong super-conditionalizing rule for c, then $\langle c, R \rangle$ is not vulnerable to a weak Dutch Strategy.*

Thus, by Theorem 3(ii), $\langle c, R_3 \rangle$ is not vulnerable even to a weak Dutch Strategy. The DSA, then, cannot say what is irrational about Condi if she begins with prior c and either endorses R_3 as an updating plan or is disposed to update in line with it. Thus, the DSA cannot justify Deterministic Updating. And without DU, it cannot support PC or DC either. After all, R_3 violates each of those, but it is not vulnerable even to a weak Dutch Strategy. And moreover, each of the three arguments for AC break down because they depend on PC or DC. The problem is that, if Condi updates from c to c_F° upon learning F, she violates AC; but there is an updating rule—namely, R_3—that allows c_F° as a response to learning F, and, for all DSA tells us, she might have rationally endorsed R_3 before learning F or she might rationally have been disposed to follow it. Indeed, the only restriction that DSA can place on your actual updating behaviour is that you should become certain of the evidence that you learned. After all:

Theorem 4 *Suppose c is your prior and c' is your posterior. Then, providing $c'(E_i) = 1$, there is a rule R such that:*

  (i) *c' is in C_i, and*

  (ii) *R is a strong super-conditionalizing rule for c.*

Thus, at the end of this section, we can conclude that, whatever is irrational about planning to update using non-deterministic updating rules that are nonetheless strong super-conditionalizing rules for your prior, it cannot be that following through on those plans leaves you vulnerable to a Dutch Strategy, for it does not. And similarly, whatever is irrational about being disposed to update in those ways, it cannot be that those dispositions will equip you with credences that lead you to choose dominated options, for they do not. With PC and DC thus blocked, our route to AC is therefore also blocked.

4 The Expected Pragmatic Utility Argument (EPUA)

Let's look at EPUA next. Again, I will consider how our credences guide our actions when we face decision problems. In this case, there is no need to restrict attention to monetary decision problems. I will only consider a single decision problem, which we face at the later time, after we've received the evidence, so I won't have to combine the outcomes of multiple options as I did in the DSA. The idea is this. Suppose you will make a decision after you receive whatever evidence it is that you receive at the later time. And suppose that you will use your later updated credence function to make that choice—indeed, you'll choose from the available options by maximising expected utility from the point of view of your new updated credences. Which updating rules does your prior expect will lead you to make the choice best?

4.1 EPUA with Deterministic Updating

Suppose you'll face decision problem $\mathbf{d}$ after you've updated. And suppose further that you'll use a deterministic updating rule R. Then, if w is a possible world and E_i is the element of the evidential partition $\mathcal{E}$ that is true at w, the idea is that we take the pragmatic utility of R relative to $\mathbf{d}$ at w to be the utility at w of whatever option from $\mathbf{d}$ we should choose if our posterior credence function were c_i, as R requires it to be at w. But of course, for many decision problems, this isn't well defined because there isn't a *unique* option in $\mathbf{d}$ that maximises expected utility by the lights of c_i; rather there are sometimes many such options, and they might have different utilities at w. Thus, we need not only c_i but also a selection function, which picks a single option from any set of options. If f is such a selection function, then let $A^{\mathbf{d}}_{c_i, f}$ be the option that f selects from the set of options in $\mathbf{d}$ that maximise expected utility by the lights of c_i. And let

$$u_{\mathbf{d}, f}(R, w) = u(A^{\mathbf{d}}_{c_i, f}, w).$$

Then the EPUA argument turns on the following mathematical fact (Savage, 1954; Good, 1967; Brown, 1976):

Theorem 5 *Suppose R and $R^{\star}$ are both deterministic updating rules. Then:*

(i) *If R and $R^{\star}$ are both conditionalizing rules for c, and f, g are selection functions, then for all decision problems $\mathbf{d}$,*

$$\sum_{w \in W} c(w) u_{\mathbf{d}, f}(R, w) = \sum_{w \in W} c(w) u_{\mathbf{d}, g}(R^{\star}, w)$$

(ii) *If R is a conditionalizing rule for c, and $R^{\star}$ is not, and f, g are selection functions, then for all decision problems $\mathbf{d}$,*

$$\sum_{w \in W} c(w) u_{\mathbf{d}, f}(R, w) \geq \sum_{w \in W} c(w) u_{\mathbf{d}, g}(R^{\star}, w)$$

with strict inequality for some decision problems $\mathbf{d}$.

That is, a deterministic updating rule maximises expected pragmatic utility by the lights of your prior just in case it is a conditionalizing rule for your prior.

As in the case of the DSA above, then, if we assume Deterministic Updating (DU), we can establish PC and DC. On the back of those, we can establish AC as well, using one of the arguments from the end of Section 3.1. After all, it is surely irrational to plan to update in one way when you expect another way to guide your actions better in the future; and it is surely irrational to be disposed to update in one way when you expect another to guide you better. And as before there are the same three arguments for AC on the back of PC and DC.

4.2 EPUA without Deterministic Updating

How does EPUA fare when we widen our view to include non-deterministic updating rules as well? The problem is that it is not clear how to define the pragmatic utility of such an updating rule relative to a decision problem and selection function at a possible world. Above, I said that, relative to a decision problem $\mathbf{d}$ and a selection function f, the pragmatic utility of rule R at world w is the utility of the option that you would choose when faced with $\mathbf{d}$ using the credence function that R mandates at w and f: that is, if E_i is true at w, then

$$u_{\mathbf{d},f}(R,w) = u(A^{\mathbf{d}}_{c_i,f}, w).$$

But, if R is not deterministic, there might be no single credence function that it mandates at w. If E_i is the piece of evidence you'll learn at w and R permits more than one credence function in response to E_i, then there might be a range of different options in $\mathbf{d}$, each of which maximises expected utility relative to a different credence function in C_i. So what are we to do?

There are (at least) two possibilities: *the fine-graining response* and *the coarse-graining response*. On the former, we cannot establish PC or DC without assuming DU; on the latter, we can.

Let's begin with the former. When we notice that there might be no single credence function that our rule R mandates at world w, a natural response is to say that I should specify our worlds in more detail, so that they determine not only the truth or falsity of the propositions in $\mathcal{F}$, but also which credence function you in fact adopt from those that R permits. In fact, given that we will be comparing the expected pragmatic utility of two different updating rules, R and $R^\star$, we need worlds that specify not only a credence function that someone following R adopts but also a credence function that someone following $R^\star$ adopts. If w is in E_i, c' is in C_i,

and $c^{\star\prime}$ is in $C_i^\star$, then let w & $R_{c\prime}^i$ & $R_{c^{\star\prime}}^{\star i}$ be the world at which the propositions in $\mathcal{F}$ that are true at w are true, the propositions in $\mathcal{F}$ that are false at w are false, the person with rule R adopts $c\prime$ in response to receiving evidence E_i and the person with $R^\star$ adopts $c^{\star\prime}$ in response to that evidence. And define the pragmatic utilities of R and $R^\star$ at this world relative to a decision problem **d** and a selection function f in the natural way:

- $u_{\mathbf{d},f}(R, w \ \& \ R_{c\prime}^i \ \& \ R_{c^{\star\prime}}^{\star i}) = u(A_{c\prime,f}^{\mathbf{d}}, w)$

- $u_{\mathbf{d},f}(R^\star, w \ \& \ R_{c\prime}^i \ \& \ R_{c^{\star\prime}}^{\star i}) = u(A_{c^{\star\prime},f}^{\mathbf{d}}, w)$

The problem, of course, is that, in the EPUA, we wish to calculate the expected pragmatic utility of an updating rule from the point of view of the prior. And that's possible only if the prior assigns a credence to each of the possible worlds. But, while our assumption that $\mathcal{F}$ is a finite algebra guarantees that a prior defined on $\mathcal{F}$ assigns a credence to each w in W, there is no guarantee that it assigns one to each w & $R_{c\prime}^i$ & $R_{c^{\star\prime}}^{\star\prime}$. So what's to be done?

A natural proposal is this: an updating rule R is rationally permissible from the point of view of a prior c just in case there is some way to extend c to c^* such that R maximises expected pragmatic utility by the lights of the extended prior, c^*. However, it is straightforward to see that any super-conditionalizing rule for a prior is rationally permissible by this standard. After all, if R is a weak or strong super-conditionalizing rule for c, then there is an extension of c^* such that R is a conditionalizing rule for c^*, and then we can piggyback on Theorem 5.

Theorem 6 *Suppose R is an updating rule. Then, if R is a weak or strong super-conditionalizing rule for c, then there is an extension c^* of c such that, for all updating rules $R^\star$, for all selection functions f, g, and all decision problems* **d**

$$\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c\prime \in C_i} \sum_{c^{\star\prime} \in C_i^\star} c^*(w \ \& \ R_{c\prime}^i \ \& \ R_{c^{\star\prime}}^{\star i}) u_{\mathbf{d},f}(R, w \ \& \ R_{c\prime}^i \ \& \ R_{c^{\star\prime}}^{\star i})$$
$$\geq \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c\prime \in C_i} \sum_{c^{\star\prime} \in C_i^\star} c^*(w \ \& \ R_{c\prime}^i \ \& \ R_{c^{\star\prime}}^{\star i}) u_{\mathbf{d},g}(R^\star, w \ \& \ R_{c\prime}^i \ \& \ R_{c^{\star\prime}}^{\star i})$$

So, if we opt for the fine-graining response to the problem of defining the pragmatic utility of a non-deterministic rule at a world, then we cannot establish either PC or DC without assuming DU and restricting the set of permissible updating rules to include only the deterministic ones.

But we might instead adopt the coarse-graining response. On this response, we retain the original possible worlds w in W, and we define the pragmatic utility of a rule at a world as either the expectation or the average of its pragmatic utility, depending on whether we are thinking of the rule as representing our dispositions or our plans, and thus whether we aim to establish DC or PC.

Suppose, first, that we are interested in DC. That is, we are interested in a norm that governs the updating rule that records how you are disposed

16

to update when you receive certain evidence. Then it seems reasonable to assume that the updating rule that records your dispositions is stochastic. That is, for each possible piece of evidence E_i and each possible response c' in C_i to that evidence that you might adopt, there is some objective chance that you will respond to E_i by adopting c'. As I explained above, I'll write this $P(R_{c'}^i | E_i)$, where $R_{c'}^i$ is the proposition that you receive E_i and respond by adopting c'. Then, if E_i is true at w, we might take the pragmatic utility of R relative to $\mathbf{d}$ and f at w to be the expectation of the utility of the options that each permitted response to E_i (and selection function f) would lead us to choose:[6]

$$u_{\mathbf{d},f}(R, w) = \sum_{c' \in C_i} P(R_{c'}^i | E_i) u(A_{c',f}^{\mathbf{d}}, w)$$

With this in hand, we have the following result:

Theorem 7 *Suppose R and $R^\star$ are both updating rules. Then:*

(i) *If R and $R^\star$ are both conditionalizing rules for c, and f, g are selection functions, then for all decision problems $\mathbf{d}$,*

$$\sum_{w \in W} c(w) u_{\mathbf{d},f}(R, w) = \sum_{w \in W} c(w) u_{\mathbf{d},g}(R^\star, w)$$

(ii) *R is a conditionalizing rule for c, and $R^\star$ is a stochastic but not conditionalizing rule, and f, g are selection functions, then for all decision problems $\mathbf{d}$,*

$$\sum_{w \in W} c(w) u_{\mathbf{d},f}(R, w) \geq \sum_{w \in W} c(w) u_{\mathbf{d},g}(R^\star, w)$$

with strictly inequality for some decision problems $\mathbf{d}$.

This shows the first difference between the DSA and EPUA. The latter, but not the former, provides a route to establishing Dispositional Conditionalization (DC). If we adopt the coarse-graining response to the problem of defining the pragmatic utility of an updating rule, we can establish DC. If we assume that your dispositions are governed by a chance function, and we use that chance function to calculate expectations, then we can show that your prior will expect your posteriors to do worse as a guide to action unless you are disposed to update by conditionalizing on the evidence you receive.

Next, suppose we are interested in Plan Conditionalization (PC). In this case, we might try to appeal again to Theorem 7. To do that, we must assume that, while there are non-deterministic updating rules that we might endorse, they are all at least stochastic updating rules; that is, they all come equipped with a probability function that determines how likely it is that I will adopt a particular permitted response to the evidence I receive.

---

[6]Recall: we assumed that each C_i is finite, so this is well-defined.

That is, we might say that the updating rules that we might endorse are either deterministic or non-deterministic-but-stochastic. In the language of game theory, we might say that the updating strategies between which we choose are either pure or mixed. And then Theorem 7 will show that we should adopt a deterministic-and-conditionalizing rule, rather than any deterministic-but-non-conditionalizing or non-deterministic-but-stochastic rule. The problem with this proposal is that it seems just as arbitrary to restrict to deterministic and non-deterministic-but-stochastic rules as it was to restrict to deterministic rules in the first place. Why should we not be able to endorse a non-deterministic and non-stochastic rule—that is, a rule that says, for at least one possible piece of evidence E_i in $\mathcal{E}$, there are two or more posteriors that the rule permits as responses, but does not endorse any chance mechanism by which we'll choose between them? But if we permit these rules, how are we to define their pragmatic utility relative to a decision problem and at a possible world?

Here's one suggestion. Suppose E_i is the proposition in $\mathcal{E}$ that is true at world w. And suppose $\mathbf{d}$ is a decision problem and f is a selection rule. Then we might take the pragmatic utility of R relative to $\mathbf{d}$ and f and at w to be the average (specifically, the mean) utility of the options that each permissible response to E_i and f would choose when faced with $\mathbf{d}$. That is,

$$u_{\mathbf{d},f}(R, w) = \frac{1}{|C_i|} \sum_{c' \in C_i} u(A^{\mathbf{d}}_{c',f}, w)$$

where $|C_i|$ is the size of C_i, that is, the number of possible responses to E_i that R permits.[7] If that's the case, then we have the following:

Theorem 8 *Suppose R and $R^\star$ are updating rules. Then if R is a conditionalizing rule for c, and $R^\star$ is not deterministic, not stochastic, and not a conditionalizing rule for c, and f, g are selection functions, then for all decision problems $\mathbf{d}$,*

$$\sum_{w \in W} c(w) u_{\mathbf{d},f}(R, w) \geq \sum_{w \in W} c(w) u_{\mathbf{d},f}(R^\star, w)$$

with strictly inequality for some decision problems $\mathbf{d}$.

Put together with Theorems 5 and 7, this shows that our prior expects us to do better by endorsing a conditionalizing rule than by endorsing any other sort of rule, whether that is a deterministic and non-conditionalizing rule, a non-deterministic but stochastic rule, or a non-deterministic and non-stochastic rule.

So, again, we see a difference between DSA and EPUA. Just as the latter, but not the former, provides a route to establishing DC without assuming Deterministic Updating, so the latter but not the former provides a route

---

[7]Again, recall that each C_i is finite, so this is well-defined.

to establishing PC without DU. And from both of those, we have the usual three routes to AC enumerated at the end of Section 3.1. This means that, if we respond to the problem of defining pragmatic utility by taking the coarse-graining appoach, the EPUA could explain what's irrational about endorsing a non-deterministic updating rule, or having dispositions that match one. If you do, there's some alternative updating rule that your prior expects to do better as a guide to future action.

5   Expected Epistemic Utility Argument (EEUA)

The previous two arguments criticized non-conditionalizing updating rules from the standpoint of pragmatic utility. The EEUA and EUDA both criticize such rules from the standpoint of epistemic utility. The idea is this: just as credences play a pragmatic role in guiding our actions, so they play other roles as well—they represent the world; they respond to evidence; they might combine more or less coherently. These roles are purely epistemic, and so just as I defined the *pragmatic* utility of a credence function at a world when faced with a decision problem, so we can also define the *epistemic* utility of a credence function at a world—it is a measure of how valuable it is to have that credence function from a purely epistemic point of view.

5.1   EEUA with Deterministic Updating

I won't give an explicit definition of the epistemic utility of a credence function at a world. Rather, I'll simply state two properties that I'll take measures of such epistemic utility to have. These are widely assumed in the literature on epistemic utility theory and accuracy-first epistemology, and I'll defer to the arguments in favour of them that are outlined there (Joyce, 2009; Pettigrew, 2016; Horowitz, 2019).

A local epistemic utility function is a function s that takes a single credence and a truth value—either true (1) or false (0)—and returns the epistemic value of having that credence in a proposition with that truth value. Thus, $s(1, p)$ is the epistemic value of having credence p in a truth, while $s(0, p)$ is the epistemic value of having credence p in a falsehood. A global epistemic utility function is a function EU that takes an entire credence function defined on $\mathcal{F}$ and a possible world and returns the epistemic value of having that credence function when the propositions in $\mathcal{F}$ have the truth values they have in that world.

> **Strict Propriety**   A local epistemic utility function s is *strictly proper* if
>
> (i) $s(1, x)$ and $s(0, x)$ are continuous functions of x;

(ii) each credence expects itself and only itself to have the greatest epistemic utility. That is, for all $0 \leq p \leq 1$,

$$ps(1, x) + (1 - p)s(0, x)$$

is uniquely maximised, as a function of x, at $x = p$.[8]

Additivity A global epistemic utility function is *additive* if, for each proposition X in $\mathcal{F}$, there is a local epistemic utility function s_X such that the epistemic utility of a credence function c at a possible world is the sum of the epistemic utilities at that world of the credences it assigns. If w is a possible world and we write $w(X)$ for the truth value (0 or 1) of proposition X at w, this says:

$$EU(c, w) = \sum_{X \in \mathcal{F}} s_X(w(X), c(X))$$

We can then define the epistemic utility of a deterministic updating rule R in the same way we defined its pragmatic utility above: if E_i is true at w, and $C_i = \{c_i\}$, then

$$EU(R, w) = EU(c_i, w)$$

Then the standard formulation of the EEUA turns on the following theorem (Greaves & Wallace, 2006):

Theorem 9 *Suppose R and $R^\star$ are deterministic updating rules. Then:*

(i) *If R and $R^\star$ are both conditionalizing rules for c, then*

$$\sum_{w \in W} c(w)EU(R, w) = \sum_{w \in W} c(w)EU(R^\star, w)$$

(ii) *If R is a conditionalizing rule for c and $R^\star$ is not, then*

$$\sum_{w \in W} c(w)EU(R, w) > \sum_{w \in W} c(w)EU(R^\star, w)$$

That is, a deterministic updating rule maximises expected epistemic utility by the lights of your prior just in case it is a conditionalizing rule for your prior.

So, as for DSA and EPUA, if we assume Deterministic Updating, we obtain an argument for PC and DC, and indirectly arguments for AC too.

---

[8]That is, if $p \neq q$, then

$$ps(1, p) + (1 - p)s(0, p) > ps(1, q) + (1 - p)s(0, q)$$

5.2 EEUA without Deterministic Updating

If we don't assume Deterministic Updating, the situation here is very similar to the one we encountered above when we considered EPUA. Suppose R is a non-deterministic updating rule. Then again we have two choices: the fine- and the coarse-graining response.

On the fine-graining response, the epistemic utility of R at a fine-grained world is

$$EU(R, w \text{ \& } R^i_{c'} \text{ \& } R^{\star i}_{c^{\star\prime}}) = EU(c', w)$$

In this case, as with EPUA, we have:

Theorem 10 *Suppose R and $R^\star$ are both updating rules. Then, if R is a weak or strong super-conditionalizing rule for c, then there is an extension c^* of c such that*

$$\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} \sum_{c^{\star\prime} \in C^\star_i} c^*(w \text{ \& } R^i_{c'} \text{ \& } R^{\star i}_{c^{\star\prime}}) EU(R, w \text{ \& } R^i_{c'} \text{ \& } R^{\star i}_{c^{\star\prime}})$$
$$\geq \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} \sum_{c^{\star\prime} \in C^\star_i} c^*(w \text{ \& } R^i_{c'} \text{ \& } R^{\star i}_{c^{\star\prime}}) EU(R^\star, w \text{ \& } R^i_{c'} \text{ \& } R^{\star i}_{c^{\star\prime}})$$

So it seems that super-conditionalizing updating rules are rationally permissible, at least by the lights of expected epistemic utility.

Next, the coarse-graining response. Suppose R is non-deterministic but stochastic. Then we let its epistemic utility at a coarse-grained world be the expectation of the epistemic utility that the various possible posteriors permitted by R take at that world. That is, if E_i is the proposition in $\mathcal{E}$ that is true at w, then

$$EU(R, w) = \sum_{c' \in C_i} P(R^i_{c'} | E_i) EU(c', w)$$

Then, we have a similar result to Theorem 7:

Theorem 11 *Suppose R and $R^\star$ are updating rules. Then if R is a conditionalizing rule for c, and $R^\star$ is stochastic but not a conditionalizing rule for c, then*

$$\sum_{w \in W} c(w) EU(R, w) > \sum_{w \in W} c(w) EU(R^\star, w)$$

Next, suppose R is a non-deterministic but also a non-stochastic rule. Then we let its epistemic utility at a world be the average epistemic utility that the various possible posteriors permitted by R take at that world. That is, if E_i is the proposition in $\mathcal{F}$ that is true at w, then

$$EU(R, w) = \frac{1}{|C_i|} \sum_{c' \in C_i} EU(c', w)$$

And again we have a similar result to Theorem 8:

21

Theorem 12 *Suppose R and R⋆ are updating rules. Then if R is a conditionalizing rule for c, and R⋆ is not deterministic, not stochastic, and not a conditionalizing rule for c. Then:*

$$\sum_{w \in W} c(w)EU(R, w) > \sum_{w \in W} c(w)EU(R^\star, w)$$

So the situation is the same as for EPUA. If we take the coarse-graining approach, whether we assess a rule by looking at how well the posteriors it produces guide our future actions or how good they are from a purely epistemic point of view, our prior will expect a conditionalizing rule for itself to be better than any non-conditionalizing rule. And thus we obtain PC and DC, and indirectly AC as well.

6 Epistemic Utility Dominance Argument (EUDA)

Finally, I turn to the EUDA. In EPUA and EEUA, we assess the pragmatic or epistemic utility of the updating rule from the viewpoint of the prior. In DSA, we assess the prior and updating rule together, and from no particular point of view; and, unlike the EPUA and EEUA, we do not assign utilities, either pragmatic or epistemic, to the prior and the rule. In EUDA, like in DSA and unlike EPUA and EEUA, we assess the prior and updating rule together, and again from no particular point of view; but, unlike in DSA and like in EPUA and EEUA, we assign utilities to them—in particular, epistemic utilities—and assess them with reference to those.

6.1 EUDA with Deterministic Updating

Suppose R is a deterministic updating rule. Then, as before, if E_i is true at w, let the epistemic utility of R be the epistemic utility of the credence function c_i that it mandates at w: that is,

$$EU(R, w) = EU(c_i, w),$$

But this time also let the epistemic utility of the pair $\langle c, R \rangle$ consisting of the prior and the updating rule be the sum of the epistemic utility of the prior and the epistemic utility of the updating rule: that is,

$$EU(\langle c, R \rangle, w) = EU(c, w) + EU(R, w) = EU(c, w) + EU(c_i, w)$$

Then the EUDA turns on the following mathematical fact (Briggs & Pettigrew, 2018):

Theorem 13 *Suppose EU is an additive, strictly proper epistemic utility function. And suppose R and R⋆ are deterministic updating rules. Then:*

(i) *If $\langle c, R \rangle$ is not conditionalizing, there is $\langle c^\star, R^\star \rangle$ such that, for all w,*

$$EU(\langle c, R \rangle, w) < EU(\langle c^\star, R^\star \rangle, w))$$

(ii) *If $\langle c, R \rangle$ is conditionalizing, there is no $\langle c^\star, R^\star \rangle$ such that, for all w,*

$$EU(\langle c, R \rangle, w) < EU(\langle c^\star, R^\star \rangle, w))$$

That is, if R is not a conditionalizing rule for c, then together they are *EU*-dominated; if it is a conditionalizing rule, they are not. Thus, like EPUA and EEUA and unlike DSA, if we assume Deterministic Updating, EUDA gives PC, DC, and indirectly AC.

6.2 EUDA without Deterministic Updating

Now suppose we permit non-deterministic updating rules as well as deterministic ones. As before, there are two approaches: the fine- and coarse-graining approaches. Here is the relevant result for the fine-graining approach:

Theorem 14 *Suppose EU is an additive, strictly proper epistemic utility function. Then:*

(i) *If R is a weak or strong super-conditionalizing rule for c, there is no $\langle c^\star, R^\star \rangle$ such that, for all E_i in $\mathcal{E}$, w in E_i, c' in C_i and $c^{\star\prime}$ in $C_i^\star$*

$$EU(\langle c, R \rangle, w \& R_{c'}^i \& R_{c^{\star\prime}}^{\star i}) < EU(\langle c^\star, R^\star \rangle, w \& R_{c'}^i \& R_{c^{\star\prime}}^{\star i})$$

(ii) *There are rules R that are not weak or strong super-conditionalizing rules for c such that there is no $\langle c^\star, R^\star \rangle$ such that, for all E_i in $\mathcal{E}$, w in E_i, c' in C_i and $c^{\star\prime}$ in $C_i^\star$*

$$EU(\langle c, R \rangle, w \& R_{c'}^i \& R_{c^{\star\prime}}^{\star i}) < EU(\langle c^\star, R^\star \rangle, w \& R_{c'}^i \& R_{c^{\star\prime}}^{\star i})$$

Interpreted in this way, then, and without the assumption of Deterministic Updating, EUDA is the weakest of all the arguments. Whereas DSA at least establishes that your updating rule should be a weak or strong super-conditionalizing for your prior, even if it does not establish that it should be conditionalizing, EUDA does not establish even that.

And here is the relevant result for the coarse-graining approach:

Theorem 15 *Suppose EU is an additive, strictly proper epistemic utility function. Then, if $\langle c, R \rangle$ is not a conditionalizing pair, there is an alternative pair $\langle c^\star, R^\star \rangle$ such that, for all w,*

$$EU(\langle c, R \rangle, w) < EU(\langle c^\star, R^\star \rangle, w)$$

This therefore supports an argument for PC and DC and indirectly AC as well.

7 Conclusion

One upshot of this investigation is that, so long as we assume Deterministic Updating (DU), all four arguments support the same conclusions, namely, Plan (PC) and Dispositional (DC) Conditionalization, and also Actual Conditionalization (AC). But once we drop DU, that agreement vanishes.

Without DU, DSA shows only that, if we plan to update using a particular rule, it should be a super-conditionalizating rule for our prior; and similarly for our dispositions. As a result, it cannot support AC. Indeed, it can support only the weakest restrictions on our actual updating behaviour, since nearly any such behaviour can be seen as an implementation of a super-conditionalizing rule—as long as we become certain of the evidence we receive after we receive it, we can be represented as having followed a strong super-conditionalizing rule.

EPUA, EEUA, and EUDA are more hopeful, at least if we adopt the coarse-graining response to the question of how to define the pragmatic or epistemic utility of a non-deterministic updating rule at a world. Let's consider our updating dispositions first. It seems natural to assume that, even if these are not deterministic, they are at least governed by objective chances. If so, and if we define the pragmatic or epistemic utility of the rule that represents those dispositions to be the expectation of the pragmatic or epistemic utility of the credence functions it produces, then we obtain DC without assuming DU. That is, we can justify DU by appealing to pragmatic or epistemic utility. However, if we use the fine-graining response, this isn't possible. And similarly when we consider our updating plans. Here, if we use the coarse-graining response and define the pragmatic or epistemic utility of the rule that represents our plan to be the average pragmatic or epistemic utility of the credence functions it produces, then we obtain PC without assuming DU. And again, if we use the fine-graining response, this isn't possible. Indeed, if we use the fine-graining approach, EPUA and EEUA establish only that you should plan to update using a weak or strong super-conditionalizing rule. And EUDA doesn't even establish that.

So, at least if we look just to our existing arguments for Conditionalization, the fates of DC and PC seem to turn on making one of two responses. First, we might simply assume Deterministic Updating. That is, to establish PC, we might simply assume that we are rationally required to plan to update in a deterministic way; and, to establish DC, we might assume that we are rationally required to have deterministic updating dispositions. This doesn't seem promising. Typically, those philosophers who offer one of the four arguments for Conditionalization studied here do so precisely because we aren't content to make brute normative assumptions about how it is rational to update; we wish to know what is so good about updating in the way that Conditionalization describes and what is so bad about

updating in some other way. Part of that is a desire to know what is so good about updating deterministically and what is so bad about updating non-deterministically. As I mentioned at the beginning, the four arguments considered here are teleological, so they specifically tell you the goods that Conditionalization and deterministic updating obtain for you. To simply assume DU is to leave the latter a mystery.

Second, we might take the coarse-graining response to the problem of defining pragmatic and epistemic utility, and then appeal to EPUA, EEUA, or EUDA. This seems more promising. That response certainly seems reasonable— that is, it produces a reasonable way to define the pragmatic or epistemic utility of an updating rule at a coarse-grained world. The problem is that we need to do more than that if we are to establish DC and PC. It is not sufficient to show that the coarse-graining response is reasonable and therefore *permissible*. We have to show that it is *mandatory*. After all, if the fine- and the coarse-graining responses are both permissible, then there is a permissible way of defining pragmatic and epistemic utility on which non-conditionalizing updating rules are permissible. And that suggests that those rules are themselves permissible. And that conflicts with DC and PC. What is needed is an argument that the fine-graining response is not legitimate. For myself, I don't see what that might be.

8 Appendix: Proofs

Recall:

(a) R is a weak super-conditionalizing rule for c if there is an extension c^* of c such that, for all E_i in $\mathcal{E}$ and c' in C_i, if $c^*(R^i_{c'}) > 0$, then $c'(-) = c^*(-|R^i_{c'})$.

(b) R is a strong super-conditionalizing rule for c if there is an extension c^* of c such that, for all E_i in $\mathcal{E}$ and c' in C_i, $c^*(R^i_{c'}) > 0$ and $c'(-) = c^*(-|R^i_{c'})$.

8.1 Dutch Strategy Argument

Lemma 2

(i) *R is a weak super-conditionalizing rule for c iff there is, for each E_i in $\mathcal{E}$ and c' in C_i, $0 \leq \lambda^i_{c'} \leq 1$ with $\sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} = 1$ such that*

    (a) *for all E_i in $\mathcal{E}$ and c' in C_i, if $\lambda^i_{c'} > 0$, then $c'(E_i) = 1$, and*

    (b) *$c(-) = \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} c'(-)$*

(ii) *R is a strong super-conditionalizing rule for c iff there is, for each E_i in $\mathcal{E}$ and c' in C_i, $0 < \lambda^i_{c'} < 1$ with $\sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} = 1$ such that*

<div align="center">25</div>

(a) *for all E_i in $\mathcal{E}$ and c' in C_i, $c'(E_i) = 1$; and*

(b) $c(-) = \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} c'(-)$

Proof of Lemma 2. Let's take (i) first. We'll begin with the left-to-right direction.

Suppose R is a weak superconditionalizing rule for c. Then, if $c^*(R^i_{c'}) > 0$, then $c'(E_i) = c^*(E_i|R^i_{c'})$. But $R^i_{c'}$ says that you received evidence E_i and responded by adopting credence function c'. So $R^i_{c'}$ entails E_i, and thus $c^*(E_i|R^i_{c'}) = 1$. So $c'(E_i) = 1$. That gives (a).

Next, for each E_i in $\mathcal{E}$ and c' in C_i, let $\lambda^i_{c'} = c^*(R^i_{c'})$. Now, for each E_i in $\mathcal{E}$ and c' in C_i, and for each possible world w, we have $c^*(w \ \& \ R^i_{c'}) = c^*(R^i_{c'})c'(w)$. Thus:

$$
\begin{aligned}
c(w) &= c^*(w) \ \text{since } c^* \text{ extends } c \\
&= \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} c^*(w \ \& \ R^i_{c'}) \ \text{ by Finite Additivity of } c^* \\
&= \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} c^*(R^i_{c'})c'(w) \ \text{ as noted above} \\
&= \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} c'(w)
\end{aligned}
$$

as required. This gives (b).

Second, we take the right-to-left direction of (i). Suppose (a) and (b) hold. Then there is, for each E_i in $\mathcal{E}$ and c' in C_i, $0 \leq \lambda^i_{c'} \leq 1$ with $\sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} = 1$ such that

$$
c(-) = \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} c'(-)
$$

So, given a possible world w, E_i in $\mathcal{E}$, and c' in C_i, let

$$
c^*(w \ \& \ R^i_{c'}) = \lambda^i_{c'} c'(w)
$$

Then

- For any possible world w,

$$
c^*(w) = \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} c^*(w \ \& \ R^i_{c'}) = \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \lambda^i_{c'} c'(w) = c(w)
$$

  So c^* is an extension of c.

- For any possible world w, E_i in $\mathcal{E}$, and c' in C_i, if $c^*(R^i_{c'}) > 0$, then

$$
c^*(w|R^i_{c'}) = \frac{c^*(w \ \& \ R^i_{c'})}{c^*(R^i_{c'})} = \frac{\lambda^i_{c'} c'(w)}{\sum_{w' \in W} \lambda^i_{c'} c'(w')} = \frac{\lambda^i_{c'} c'(w)}{\lambda^i_{c'} \sum_{w' \in W} c'(w')} = c'(w)
$$

  and thus $c'(E_i) = c^*(E_i|R^i_{c'}) = 1$.

26

Thus, R is a weak super-conditionalizing rule for c. This establishes Lemma 2(i).

The proof of Lemma 2(ii) proceeds in exactly the same way. □

Theorem 3

(i) *If R is not a weak or strong super-conditionalizing rule for c, then it is vulnerable at least to a weak Dutch Strategy, and possibly to a strong Dutch Strategy.*

(ii) *If R is a strong super-conditionalizing rule for c, then it is not vulnerable even to a weak Dutch Strategy.*

Proof of Theorem 3. First, (i). Suppose R is not a weak or strong super-conditionalizing rule for c. Then, by Lemma 2, either

(a) $c'(E_i) < 1$ for some E_i in $\mathcal{E}$ and c' in C_i; or

(b) c is not in the convex hull of R.[9]

Let's take these in turn.

First, (a). Suppose that $c'(E_i) < p < 1$ for some E_i in $\mathcal{E}$ and c' in C_i. Then let $X^i_{c'}$ be an option that has utility $-(1-p)$ if E_i is true and p if not. And let $Y^i_{c'}$ be the option that has utility 0 regardless. Then offer no decision problem at the earlier time and offer one at the later time only if the agent learns E_i and adopts c'; and in that situation, offer $\mathbf{d}^i_{c'} = \{X^i_{c'}, Y^i_{c'}\}$. Then the agent will choose $X^i_{c'}$, but that will do worse than $Y^i_{c'}$ at all worlds at which E_i is true. So $\langle c, R \rangle$ is vulnerable to a weak Dutch Strategy.

Second, (b). Suppose c is not in the convex hull of the set of posteriors that R permits. That is, c is not in the convex hull of the set $\{c' : (\exists E_i \in \mathcal{E})(c' \in C_i)\}$. Now, let's represent each credence function on $\mathcal{F}$ by the vector of the values that it takes the possible worlds w in W. Thus, if $W = \{w_1, \ldots, w_m\}$, then we represent c by $\langle c(w_1), \ldots, c(w_m) \rangle$; and, for any E_i in $\mathcal{E}$ and c' in C_i, we represent c' by $\langle c'(w_1), \ldots, c'(w_m) \rangle$. Now, since c is outside the convex hull of the set of posteriors that R permits, the vector that represents c is outside the convex hull of the set of vectors that represent the posteriors that R permits. Thus, by the Separating Hyperplane Theorem, there is a vector $S = \langle S_1, \ldots, S_m \rangle$ such that, for any E_i in $\mathcal{E}$ and c' in C_i,

$$c' \cdot S < c \cdot S$$

Then pick m, ε such that $c' \cdot S < m - \varepsilon < m < c \cdot S$. Now let:

---

[9]The convex hull of a set of points is the smallest convex set that contains it as a subset. A set is convex if, for any two points in it, any convex combination of those two points also lies in the set.

- A be the option that has utility S_i at world w_i;

- B be the option that has utility m at world w_i;

- $B - \varepsilon$ be the option that has utility $m - \varepsilon$ at world w_i.

Then, let $\mathbf{d} = \{A, B\}$. Your prior c will choose A, since the expected utility of A is $c \cdot S$, while the expected utility of B is m. And let $\mathbf{d}' = \{A, B - \varepsilon\}$. Then each of your possible posteriors c' will choose $B - \varepsilon$, since the expected utility of A is $c' \cdot S$, while the expected utility of $B - \varepsilon$ is $m - \varepsilon$. Choosing B from $\mathbf{d}$ and A from $\mathbf{d}'$ is guaranteed to have greater total utility than choosing A from $\mathbf{d}$ and $B - \varepsilon$ from $\mathbf{d}'$. So $\langle c, R \rangle$ is vulnerable to a strong Dutch Strategy. This establishes Theorem 3(i).

Second, (ii). Suppose there are decision problems $\mathbf{d}$ and $\mathbf{d}^i_{c'}$ for each E_i in $\mathcal{E}$ and c' in C_i. And suppose $\mathbf{d} = \{A, B\}$ and $\mathbf{d}^i_{c'} = \{X^i_{c'}, Y^i_{c'}\}$. Now suppose c would choose A over B and, for each E_i in $\mathcal{E}$ and c' in C_i, c' would choose $X^i_{c'}$ over $Y^i_{c'}$. So:

(a) $\sum_{w \in W} c(w) u(A, w) > \sum_{w \in W} c(w) u(B, w)$

(b) $\sum_{w \in W} c'(w) u(X^i_{c'}, w) > \sum_{w \in W} c'(w) u(Y^i_{c'}, w)$, for all E_i in $\mathcal{E}$ and c' in C_i

Now, suppose that, for each E_i in $\mathcal{E}$, w in E_i, and c' in C_i,

$$u(A, w) + u(X^i_{c'}, w) < u(B, w) + u(Y^i_{c'}, w) \tag{$\dagger$}$$

Then

$$
\begin{aligned}
&\sum_{w \in W} c(w) u(A, w) + \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} c^*(R^i_{c'}) \sum_{w \in W} c'(w) u(X^i_{c'}, w) \\
=\ &\sum_{w \in W} c^*(w) u(A, w) + \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(R^i_{c'}) c'(w) u(X^i_{c'}, w) \\
=\ &\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(w \,\&\, R^i_{c'}) u(A, w) + \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(w \,\&\, R^i_{c'}) u(X^i_{c'}, w) \\
=\ &\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(w \,\&\, R^i_{c'}) [u(A, w) + u(X^i_{c'}, w)] \\
<\ &\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(w \,\&\, R^i_{c'}) [u(B, w) + u(Y^i_{c'}, w)] \quad \text{by } (\dagger) \\
=\ &\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(w \,\&\, R^i_{c'}) u(B, w) + \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(w \,\&\, R^i_{c'}) u(Y^i_{c'}, w) \\
=\ &\sum_{w \in W} c^*(w) u(B, w) + \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} c^*(R^i_{c'}) c'(w) u(Y^i_{c'}, w) \\
=\ &\sum_{w \in W} c(w) u(B, w) + \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} c^*(R^i_{c'}) \sum_{w \in W} c'(w) u(Y^i_{c'}, w)
\end{aligned}
$$

But this contradicts (a) and (b). This establishes Theorem 3(ii). $\qquad \square$

8.2 Expected Pragmatic Utility Argument

We first prove the following lemma. Theorems 5, 7, and 8 all follow as corollaries.

Lemma 16

(i) *If R, $R^\star$ are conditionalizing rules for c, and f, g are selection functions, then for all decision problems $\mathbf{d}$*

$$\sum_{w \in W} c(w) u_{\mathbf{d},f}(R, w) = \sum_{w \in W} c(w) u_{\mathbf{d},g}(R^\star, w)$$

(ii) *If R is a conditionalizing rule for c, and $R^\star$ is not, and f, g are selection functions, then for all decision problems $\mathbf{d}$,*

$$\sum_{w \in W} c(w) u_{\mathbf{d},f}(R, w) \geq \sum_{w \in W} c(w) u_{\mathbf{d},g}(R^\star, w)$$

with strict inequality for some decision problems $\mathbf{d}$.

Proof. First, (i). Suppose R and $R^\star$ are conditionalizing rules for c, and f, g are selection functions. So:

- $R = (\mathcal{E} = \{E_1, \ldots, E_n\}, \mathcal{C} = \{C_1, \ldots, C_n\})$ and
- $R^\star = (\mathcal{E} = \{E_1, \ldots, E_n\}, \mathcal{C} = \{C_1^\star, \ldots, C_n^\star\})$.

And, if $c(E_i) > 0$,

- $C_i = \{c_i\}$ and $C_i^\star = \{c_i^\star\}$;
- $c_i(-)c(E_i) = c(- \,\&\, E_i) = c_i^\star(-)c(E_i)$;
- $c_i(-) = c(-|E_i) = c_i^\star(-)$.

Suppose $\mathbf{d}$ is a decision problem. Then,

- If $c(E_i) > 0$, then $c_i(-) = c_i^\star(-)$, and thus

$$\sum_{w \in W} c_i(w) u(A_{c_i,f}^{\mathbf{d}}, w) = \sum_{w \in W} c_i^\star(w) u(A_{c_i^\star,g}^{\mathbf{d}}, w)$$

  so

$$c(E_i) \sum_{w \in W} c_i(w) u(A_{c_i,f}^{\mathbf{d}}, w) = c(E_i) \sum_{w \in W} c_i^\star(w) u(A_{c_i^\star,g}^{\mathbf{d}}, w)$$

- If $c(E_i) = 0$, then

$$c(E_i) \sum_{w \in W} c_i(w) u(A_{c_i,f}^{\mathbf{d}}, w) = 0 = c(E_i) \sum_{w \in W} c_i^\star(w) u(A_{c_i^\star,g}^{\mathbf{d}}, w)$$

So

$$\sum_{w \in W} c(w) u_{\mathbf{d},f}(R, w)$$

$$= \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w) u(A^{\mathbf{d}}_{c_i,f}, w)$$

$$= \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(E_i) c_i(w) u(A^{\mathbf{d}}_{c_i,f}, w)$$

$$= \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in W} c_i(w) u(A^{\mathbf{d}}_{c_i,f}, w)$$

$$= \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in W} c_i^\star(w) u(A^{\mathbf{d}}_{c_i^\star,g}, w)$$

$$= \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(E_i) c_i^\star(w) u(A^{\mathbf{d}}_{c_i^\star,g}, w)$$

$$= \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w) u(A^{\mathbf{d}}_{c_i^\star,g}, w)$$

$$= \sum_{w \in W} c(w) u_{\mathbf{d},g}(R^\star, w)$$

as required. This establishes Lemma 16(i).

Second, (ii). Suppose R is a conditionalizing rule, and $R^\star$ is not, and f, g are selection functions. Then, if $c(E_i) > 0$, then $C_i = \{c_i\}$, $c_i(E_i) = 1$, and $c(w) = c_i(w)c(E_i)$. Now, suppose that, for each E_i in $\mathcal{E}$ and c' in $C_i^\star$, there is $\alpha_{c'}^i > 0$ such that for each E_i in $\mathcal{E}$, $\sum_{c' \in C_i^\star} \alpha_{c'}^i = 1$ and

$$u_{\mathbf{d},f}(R^\star, w) = \sum_{c' \in C_i^\star} \alpha_{c'}^i u(A^{\mathbf{d}}_{c',f}, w)$$

Then, for all E_i in $\mathcal{E}$ and c' in $C_i^\star$,

$$\sum_{w \in W} c_i(w) u(A^{\mathbf{d}}_{c_i,f}, w) \geq \sum_{w \in W} c_i(w) u(A^{\mathbf{d}}_{c',g}, w)$$

And thus

$$\sum_{w \in W} c(w) u_{\mathbf{d},f}(R, w)$$

$$= \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w) u(A^{\mathbf{d}}_{c_i,f}, w)$$

$$= \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in E_i} c_i(w) u(A^{\mathbf{d}}_{c_i,f}, w)$$

$$= \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in W} c_i(w) \sum_{c' \in C_i^\star} \alpha^i_{c'} u(A^{\mathbf{d}}_{c_i,f}, w)$$

$$= \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{c' \in C_i^\star} \alpha^i_{c'} \sum_{w \in W} c_i(w) u(A^{\mathbf{d}}_{c_i,f}, w)$$

$$\geq \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{c' \in C_i^\star} \alpha^i_{c'} \sum_{w \in W} c_i(w) u(A^{\mathbf{d}}_{c',g}, w)$$

$$= \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in E_i} c_i(w) \sum_{c' \in C_i^\star} \alpha^i_{c'} u(A^{\mathbf{d}}_{c',g}, w)$$

$$= \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w) \sum_{c' \in C_i^\star} \alpha^i_{c'} u(A^{\mathbf{d}}_{c',g}, w)$$

$$= \sum_{w \in W} c(w) u_{\mathbf{d},g}(R^\star, w)$$

Now, since $R^\star$ is not a conditionalizing rule for c, there is $c(E_i) > 0$ and c' in $C_i^\star$ such that $c'(-) \neq c_i(-)$. Then there is a decision problem $\mathbf{d}$ such that $A^{\mathbf{d}}_{c',g}$ does not maximise expected utility with respect to c_i. And thus

$$\sum_{w \in W} c_i(w) u(A^{\mathbf{d}}_{c_i,f}, w) > \sum_{w \in W} c_i(w) u(A^{\mathbf{d}}_{c',g}, w)$$

In this case, the inequality above is strict, as required.

Now:

- If $R^\star$ is deterministic but not conditionalizing, let $\alpha^i_{c'} = 1$ for all c' in C_i. This gives Theorem 5.

- If $R^\star$ is non-deterministic but stochastic, let $\alpha^i_{c'} = P(R^{\star i}_{c'} | E_i)$ be the probability of $R^{\star i}_{c'}$ given E_i. This gives Theorem 7.

- If $R^\star$ is non-deterministic and non-stochastic, let $\alpha^i_{c'} = \frac{1}{|C_i^\star|}$. This gives Theorem 8.

This establishes Lemma 16(ii). $\qquad\square$

8.3 Expected Epistemic Utility Argument

As in the previous section, we first prove a lemma. Theorems 9, 11, and 12 all follow as corollaries.

Lemma 17

(i) *If R, $R^\star$ are conditionalizing rules for c, and f, g are selection functions, then*

$$\sum_{w \in W} c(w) EU(R, w) = \sum_{w \in W} c(w) EU(R^\star, w)$$

(ii) *If R is a conditionalizing rule for c, and $R^\star$ is not, and f, g are selection functions, then*

$$\sum_{w \in W} c(w) EU(R, w) > \sum_{w \in W} c(w) EU(R^\star, w)$$

Proof of Lemma 17. First, (i). Suppose R and $R^\star$ are conditionalizing rules for c. Then,

- If $c(E_i) > 0$, then $c_i(-) = c_i^\star(-)$, so $EU(c_i, w) = EU(c_i^\star, w)$ and

$$c(E_i) EU(c_i, w) = c(E_i) EU(c_i^\star, w)$$

- If $c(E_i) = 0$, then

$$c(E_i) EU(c_i, w) = c(E_i) EU(c_i^\star, w)$$

So

$$
\begin{aligned}
&\sum_{w \in W} c(w) EU(R, w) \\
={}& \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w) EU(c_i, w) \\
={}& \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in W} c_i(w) EU(c_i, w) \\
={}& \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in W} c_i^\star(w) EU(c_i^\star, w) \\
={}& \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w) EU(c_i^\star, w) \\
={}& \sum_{w \in W} c(w) EU(R^\star, w)
\end{aligned}
$$

as required. This establishes Lemma 17(i).

Second, (ii). Suppose R is a conditionalizing rule, and $R^\star$ is not. Now, suppose that, for each E_i in $\mathcal{E}$ and c' in $C_i^\star$, there is $\alpha_{c'}^i > 0$ such that, for each E_i in $\mathcal{E}$, $\sum_{c' \in C_i^\star} \alpha_{c'}^i = 1$ and

$$EU(R^\star, w) = \sum_{c' \in C_i^\star} \alpha_{c'}^i EU(c', w)$$

Then, for all E_i in $\mathcal{E}$ and c' in $C_i^\star$,

$$\sum_{w \in W} c_i(w)EU(c_i, w) \geq \sum_{w \in W} c_i(w)EU(c', w)$$

with strict inequality if $c' \neq c_i$. Then

$$
\begin{aligned}
&\sum_{w \in W} c(w)EU(R, w) \\
=\ &\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w)EU(c_i, w) \\
=\ &\sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in E_i} c_i(w)EU(c_i, w) \\
=\ &\sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in W} c_i(w) \sum_{c' \in C_i^\star} \alpha_{c'}^i EU(c_i, w) \\
=\ &\sum_{E_i \in \mathcal{E}} c(E_i) \sum_{c' \in C_i^\star} \alpha_{c'}^i \sum_{w \in W} c_i(w)EU(c_i, w) \\
\geq\ &\sum_{E_i \in \mathcal{E}} c(E_i) \sum_{c' \in C_i^\star} \alpha_{c'}^i \sum_{w \in W} c_i(w)EU(c', w) \\
=\ &\sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in E_i} c_i(w) \sum_{c' \in C_i^\star} \alpha_{c'}^i EU(c', w) \\
=\ &\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} c(w) \sum_{c' \in C_i^\star} \alpha_{c'}^i EU(c', w) \\
=\ &\sum_{w \in W} c(w)EU(R^\star, w)
\end{aligned}
$$

Now, since $R^\star$ is not a conditionalizing rule for c, there is $c(E_i) > 0$ and c' in $C_i^\star$ such that $c' \neq c_i$. Thus

$$\sum_{w \in W} c_i(w)EU(c_i, w) > \sum_{w \in W} c_i(w)u(c', w)$$

In this case, the inequality above is strict, as required. Now:

- If $R^\star$ is deterministic but not conditionalizing, let $\alpha_{c'}^i = 1$ for all c' in $C_i^\star$. This gives Theorem 9.

- If $R^\star$ is non-deterministic but stochastic, let $\alpha_{c'}^i = P(R_{c'}^{\star i} | E_i)$ be the probability of $R_{c'}^{\star i}$ given E_i. This gives Theorem 11.

- If $R^\star$ is non-deterministic and non-stochastic, let $\alpha_{c'}^i = \frac{1}{|C_i^\star|}$. This gives Theorem 12.

This establishes Lemma 16(ii). $\qquad\square$

9 Accuracy dominance argument

Theorem 13 *Suppose EU is an additive, strictly proper epistemic utility function. And suppose c is a prior and R is a deterministic updating rule. Then:*

(i) *if $\langle c, R \rangle$ is non-conditionalizing, there is $\langle c^\star, R^\star \rangle$ such that, for all w*

$$EU(\langle c, R \rangle, w) < EU(\langle c^\star, R^\star \rangle, w))$$

(ii) *if $\langle c, R \rangle$ is conditionalizing, then for any $\langle c^\star, R^\star \rangle$ there is some w such that*

$$EU(\langle c, R \rangle, w) \geq EU(\langle c^\star, R^\star \rangle, w))$$

Proof of Theorem 13. First, (i). Suppose R is deterministic but not conditionalizing for c. This is the case covered by (Briggs & Pettigrew, 2018). So $R = (\mathcal{E} = \{E_1, \ldots, E_n\}, \mathcal{C} = \{C_1, \ldots, C_n\})$, and, for each E_i, $C_i = \{c_i\}$. So we can write $\langle c, R \rangle$ as a $(n+1)$-dimensional vector of credence functions:

$$\langle c, c_1, \ldots, c_n \rangle$$

Now consider the following set of updating plans:

$$W_R = \{w_R = \langle w, c_1, \ldots, c_{i-1}, w, c_{i+1}, \ldots, c_n \rangle : E_i \in \mathcal{E} \ \& \ w \in E_i\}$$

Thus, w_R is the updating plan that has the credence function w as its prior, and will update exactly as R does except in world w where it will stick with w. Then we can show that $\langle c, R \rangle$ is not in the convex hull of W_R. After all:

Lemma 18 *$\langle c, R \rangle$ is in the convex hull of W_R iff R is conditionalizing for c.*

Proof of Lemma 18. Let's prove the left-to-right direction first. Suppose $\langle c, R \rangle$ is in the convex hull of W_R. Then there is $0 \leq \alpha_w \leq 1$ such that

- $c(-) = \sum_{w \in W} \alpha_w w(-)$;

- $c_i(-) = \sum_{w \in E_i} \alpha_w w(-) + \sum_{w \notin E_i} \alpha_w c_i(-)$

Thus,

$$c_i(X) = c(XE_i) + c(\overline{E_i})c_i(X)$$

And so

$$c_i(X) = c(X|E_i)$$

as required.

Next, the right-to-left direction. Suppose R is conditionalizing for c. So $c_i(X) = c(X|E_i)$. Then let $\alpha_w = c(w)$. It is easy to check that $\langle c, R \rangle$ is in the convex hull of W_R, as required. □

Proof of Theorem 13(i) continued. Now, suppose EU is an additive, strictly proper epistemic utility function. Then, by Proposition 2 in (Predd et al.,

2009), there is a Bregman divergence $\mathfrak{D}$ such that $EU(c,w) = -\mathfrak{D}(w,c)$. Let $\mathfrak{D}^{n+1}$ be the Bregman divergence between $(n+1)$-dimensional vectors of credence functions that sums the Bregman divergences between the $n+1$ credence functions. Thus, by Proposition 3 in (Predd et al., 2009), since $\langle c, R \rangle$ is not in the convex hull of W_R, there is $\langle c^\star, R^\star \rangle$ in the convex hull of W_R such that, for all E_i in $\mathcal{E}$ and w in E_i,

$$\mathfrak{D}^{n+1}((w,c_1,\ldots,c_{i-1},w,c_{i+1},\ldots,c_n),(c^\star,c_1^\star,\ldots,c_n^\star)) <$$
$$\mathfrak{D}^{n+1}((w,c_1,\ldots,c_{i-1},w,c_{i+1},\ldots,c_n),(c,c_1,\ldots,c_n))$$

But

$$\mathfrak{D}^{n+1}((w,c_1,\ldots,c_{i-1},w,c_{i+1},\ldots,c_n),(c,c_1,\ldots,c_n)) =$$
$$\mathfrak{D}(w,c) + \mathfrak{D}(w,c_i) = -EU(R,w)$$

And

$$\mathfrak{D}^{n+1}((w,c_1,\ldots,c_{i-1},w,c_{i+1},\ldots,c_n),(c^\star,c_1^\star,\ldots,c_n^\star)) \geq$$
$$\mathfrak{D}(w,c^\star) + \mathfrak{D}(w,c_i^\star) = -EU(R^\star,w)$$

So $EU(R^\star,w) > EU(R,w)$ for all w in W, as required. This establishes Theorem 13(i).

Second, (ii). Suppose $\langle c, R \rangle$ is conditionalizing, and suppose $\langle c^\star, R^\star \rangle$ is an alternative pair where $R^\star$ is deterministic. Then we show that c expects $\langle c, R \rangle$ to have at least as much epistemic utility as it expects $\langle c^\star, R^\star \rangle$ to have. So the latter cannot EU-dominate the former.

$$\sum_{w \in W} c(w) EU(\langle c, R \rangle, w)$$
$$= \sum_{w \in W} c(w) EU(c,w) + \sum_{w \in W} c(w) EU(R,w)$$
$$= \sum_{w \in W} c(w) EU(c,w) + \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in E_i} c_i(w) EU(c_i,w)$$
$$\geq \sum_{w \in W} c(w) EU(c^\star,w) + \sum_{E_i \in \mathcal{E}} c(E_i) \sum_{w \in E_i} c_i(w) EU(c_i^\star,w)$$
$$= \sum_{w \in W} c(w) EU(c^\star,w) + \sum_{w \in W} c(w) EU(R^\star,w)$$
$$= \sum_{w \in W} c(w) EU(\langle c^\star, R^\star \rangle, w)$$

as required. This establishes Theorem 13(ii). □

Before we prove our next theorem, we state and prove this lemma:

Lemma 19 *Suppose EU is an additive, strictly proper epistemic utility function. Suppose $C = \{c_1, \ldots, c_m\}$ is a finite set of credence functions where $m > 1$. Suppose $\alpha_1, \ldots, \alpha_m > 0$ and $\sum_{j=1}^m \alpha_j = 1$. Then there is a credence function $c^\star$ such that, for w in W,*

$$EU(c^\star,w) > \sum_{j=1}^m \alpha_j EU(c,w)$$

35

Proof of Lemma 19. Suppose *EU* is an additive, strictly proper epistemic utility function. Then, for each possible world w, we abuse notation and write w for the credence function defined on $\mathcal{F}$ such that $w(X) = 1$ if X is true at w and $w(X) = 0$ if X is false at w. Then, by Proposition 2 from (Predd et al., 2009), there is a Bregman divergence $\mathfrak{D}$ that measures the divergence from one credence function to another such that $EU(c, w) = -\mathfrak{D}(w, c)$.

Suppose C contains $m > 1$ credence functions. And suppose $\alpha_1, \ldots, \alpha_m > 0$ and $\sum_{j=1}^{m} \alpha_j = 1$. We can then use $\mathfrak{D}$ to generate a Bregman divergence between two m-tuples of credence functions as follows:

$$\mathfrak{D}^\alpha((c_1, \ldots, c_m), (c_1', \ldots, c_m')) = \sum_{j=1}^{m} \alpha_j \mathfrak{D}(c_j, c_j')$$

So

$$\mathfrak{D}^\alpha((w, \ldots, w), (c_1, \ldots, c_m)) = -\sum_{j=1}^{m} \alpha_j EU(c_j, w)$$

Now consider the following set of m-tuples of credence functions:

$$\mathcal{W} = \{(w, \ldots, w) : w \in W\}$$

Then $(c_1, \ldots c_m)$ is in the convex hull of $\mathcal{W}$ iff $c_1 = c_2 = \ldots = c_m$. Thus, by Proposition 3 in (Predd et al., 2009), if $c_i \neq c_j$ for some $1 \leq i, j \leq m$, then there is $(c^\star, \ldots, c^\star)$ in the convex hull of $\mathcal{W}$ such that, for w in W,

$$\mathfrak{D}^\alpha((w, \ldots, w), (c^\star, \ldots, c^\star)) < \mathfrak{D}^\alpha((w, \ldots, w), (c_1, \ldots, c_m))$$

So

$$\sum_{j=1}^{m} \alpha_j EU(c^\star, w) = EU(c^\star, w) > \sum_{j=1}^{m} \alpha_j EU(c_j, w)$$

as required. $\qquad\square$

Theorem 15 *Suppose EU is an additive, strictly proper epistemic utility function. Then, if $\langle c, R \rangle$ is non-conditionalizing, there is $\langle c^\star, R^\star \rangle$ such that, for all w,*

$$EU(\langle c, R \rangle, w) < EU(\langle c^\star, R^\star \rangle, w)$$

Proof of Theorem 15. Suppose $\langle c, R \rangle$ is non-conditionalizing. If R is deterministic, we can appeal to Theorem 13. If R is non-deterministic, then we can use Lemma 19 to construct a dominating pair $\langle c^\star, R^\star \rangle$. After all, if C_i contains more than one possible posterior, then there are $\alpha_{c'}^i > 0$ such that $\sum_{c' \in C_i} \alpha_{c'}^i = 1$ and, for w in E_i,

$$EU(R, w) = \sum_{c' \in C_i} \alpha_{c'}^i EU(c', w)$$

If R is stochastic, then $\alpha^i_{c'} = P(R^i_{c'}|E_i)$; if R is non-stochastic, then $\alpha^i_{c'} = \frac{1}{|C_i|}$. Now, by Lemma 19, there is $c^\star_i$ such that, for all w in W,

$$EU(c^\star_i, w) > \sum_{c' \in C_i} \alpha^i_{c'} EU(c', w)$$

Thus, let $R^\star$ be the updating rule such that $C^\star_j = C_j$ for all $j \neq i$, and $C^\star_i = \{c^\star_i\}$. Also, let $c^\star = c$. Then:

$$EU(\langle c^\star, R^\star \rangle, w) > EU(\langle c, R \rangle, w)$$

as required. □

Theorem 14 *Suppose EU is an additive, strictly proper epistemic utility function. Then:*

(i) *If R is a weak or strong super-conditionalizing rule for c, there is no $\langle c^\star, R^\star \rangle$ such that, for all E_i in $\mathcal{E}$, w in E_i, c' in C_i and $c^{\star\prime}$ in $C^\star_i$*

$$EU(\langle c, R \rangle, w \mathbin{\&} R^i_{c'} \mathbin{\&} R^{\star i}_{c^{\star\prime}}) < EU(\langle c^\star, R^\star \rangle, w \mathbin{\&} R^i_{c'} \mathbin{\&} R^{\star i}_{c^{\star\prime}})$$

(ii) *There are rules R that are not weak or strong super-conditionalizing rules for c such that there is no $\langle c^\star, R^\star \rangle$ such that, for all E_i in $\mathcal{E}$, w in E_i, c' in C_i and $c^{\star\prime}$ in $C^\star_i$*

$$EU(\langle c, R \rangle, w \mathbin{\&} R^i_{c'} \mathbin{\&} R^{\star i}_{c^{\star\prime}}) < EU(\langle c^\star, R^\star \rangle, w \mathbin{\&} R^i_{c'} \mathbin{\&} R^{\star i}_{c^{\star\prime}})$$

Proof of Theorem 14. First, (i). Suppose $\langle c, R \rangle$ is a weak or a strong super-conditionalizing pair. Thus, we can extend c to c^* so that, for E_i in $\mathcal{E}$ and c' in C_i, if $c^*(E_i) > 0$.
$$c'(-) = c^*(-|R^i_{c'})$$

Thus,
$$c^*(w \mathbin{\&} R^i_{c'}) = c^*(R^i_{c'})c'(w)$$

Now suppose $\langle c^\star, R^\star \rangle$ is an alternative prior-rule pair. Then extend c^* further so that:

$$c^*(w \mathbin{\&} R^i_{c'} \mathbin{\&} R^{\star i}_{c^{\star\prime}}) = \frac{1}{|C^\star_i|} c^*(w \mathbin{\&} R^i_{c'}) = \frac{1}{|C^\star_i|} c^*(R^i_{c'})c'(w)$$

Then:

$$\sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} \sum_{c^{*'} \in C_i} c^*(w \,\&\, R^i_{c'} \,\&\, R^{*i}_{c^{*'}}) EU(\langle c, R \rangle, w \,\&\, R^i_{c'} \,\&\, R^{*i}_{c^{*'}})$$

$$= \sum_{w \in W} c(w) EU(c, w) + \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \sum_{w \in E_i} c^*(w \,\&\, R^i_{c'}) EU(c', w)$$

$$= \sum_{w \in W} c(w) EU(c, w) + \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} c^*(R^i_{c'}) \sum_{w \in W} c'(w) EU(c', w)$$

$$\geq \sum_{w \in W} c(w) EU(c^\star, w) + \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} c^*(R^i_{c'}) \sum_{w \in W} c'(w) EU(c^{*'}, w)$$

$$= \sum_{w \in W} c(w) EU(c^\star, w) + \sum_{E_i \in \mathcal{E}} \sum_{c' \in C_i} \sum_{w \in W} c^*(w \,\&\, R^i_{c'}) EU(c^{*'}, w)$$

$$= \sum_{E_i \in \mathcal{E}} \sum_{w \in E_i} \sum_{c' \in C_i} \sum_{c^{*'} \in C_i} c^*(w \,\&\, R^i_{c'} \,\&\, R^{*i}_{c^{*'}}) EU(\langle c^\star, R^\star \rangle, w \,\&\, R^i_{c'} \,\&\, R^{*i}_{c^{*'}})$$

Thus, it cannot be the case that for all E_i in $\mathcal{E}$, w in E_i, c' in C_i and $c^{*'}$ in $C_i^\star$

$$EU(\langle c, R \rangle, w \,\&\, R^i_{c'} \,\&\, R^{*i}_{c^{*'}}) < EU(\langle c^\star, R^\star \rangle, w \,\&\, R^i_{c'} \,\&\, R^{*i}_{c^{*'}})$$

This establishes Theorem 14(i).

Next, (ii). Suppose $\mathcal{E} = \{E_1, \dots, E_n\}$ and w_1, w_2 are in E_1. Then pick c, c_1, c_2 so that they lie on the line between w_1 and w_2. Pick: c very close to w_1; c_1 very close to c, but slightly further towards w_2; and c_2 right at the end of the line at w_2. So c is not in the convex hull of c_1 and c_2. Next, define $R = (\mathcal{E}, \mathcal{C})$ with $C_1 = \{c_1, c_2\}$. Now suppose $c^\star$ is an alternative prior and $R^\star = (\mathcal{E}, \mathcal{C}^\star)$ is an alternative updating rule, with $c^{*'}$ in $C_1^\star$. Then, if $\langle c^\star, R^\star \rangle$ dominates $\langle c, R \rangle$, then:

(i) $EU(c, w_1) + EU(c_1, w_1) < EU(c^\star, w_1) + EU(c^{*'}, w_1)$

(ii) $EU(c, w_2) + EU(c_2, w_2) < EU(c^\star, w_2) + EU(c^{*'}, w_2)$

Now, suppose $c^\star$ is equal to c or lies between c and w_1. Then $EU(c^\star, w_2) \leq EU(c, w_2)$. But since $c_2 = w_2$, $EU(c^{*'}, w_2) \leq EU(c_2, w_2)$. So

$$EU(c, w_2) + EU(c_2, w_2) \geq EU(c^\star, w_2) + EU(c^{*'}, w_2)$$

which contradicts (ii). And similarly for $c^{*'}$. So both $c^\star$ and $c^{*'}$ must lie strictly between c and w_2. And indeed, they must lie strictly between c and c_1. If one or other lies between c_1 and w_2, then

$$EU(c, w_2) + EU(c_1, w_1) \geq EU(c^\star, w_1) + EU(c^{*'}, w_1)$$

which contradicts (i). Now, since EU is continuous in its first argument and $EU(c, w_2) < EU(c_2, w_2)$, we can always pick c_1 so that it's close enough to c that $EU(c_1, w_2)$ is close enough to $EU(c, w_2)$ that

$$EU(c_1, w_2) + EU(c_1, w_2) < EU(c, w_2) + EU(c_2, w_2)$$

38

But then, since $c^\star$ and $c^{\star\prime}$ lie between c_1 and c, we have

$$EU(c^\star, w_2) + EU(c^{\star\prime}, w_2) < EU(c_1, w_2) + EU(c_1, w_2) < EU(c, w_2) + EU(c_2, w_2)$$

which contradicts (ii). This gives our contradiction. So $\langle c^\star, R^\star \rangle$ does not dominate $\langle c, R \rangle$. □

References

Briggs, R. A., & Pettigrew, R. (2018). An accuracy-dominance argument for conditionalization. *Noûs*.

Bronfman, A. (2014). Conditionalization and Not Knowing That One Knows. *Erkenntnis*, *79*(4), 871–892.

Brown, P. M. (1976). Conditionalization and expected utility. *Philosophy of Science*, *43*(3), 415–419.

de Finetti, B. (1974). *Theory of Probability*, vol. I. New York: John Wiley & Sons.

Diaconis, P., & Zabell, S. L. (1982). Updating Subjective Probability. *Journal of the American Statistical Association*, *77*(380), 822–830.

Dietrich, F., List, C., & Bradley, R. (2016). Belief Revision generalized: A joint characterization of Bayes's and Jeffrey's rules. *Journal of Economic Theory*, *162*, 352–371.

Good, I. J. (1967). On the Principle of Total Evidence. *The British Journal for the Philosophy of Science*, *17*, 319–322.

Greaves, H., & Wallace, D. (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, *115*(459), 607–632.

Grove, A. J., & Halpern, J. Y. (1998). Updating Sets of Probabilities. In *Proceedings of the 14$^{\text{th}}$ Conference on Uncertainty in AI*, (pp. 173–182). San Francisco, CA: Morgan Kaufman.

Horowitz, S. (2019). Accuracy and Educated Guesses. In T. S. Gendler, & J. Hawthorne (Eds.) *Oxford Studies in Epistemology*, vol. 6. Oxford University Press.

Jeffrey, R. (1992). *Probability and the Art of Judgment*. New York: Cambridge University Press.

Joyce, J. M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In F. Huber, & C. Schmidt-Petri (Eds.) *Degrees of Belief*. Springer.

Konek, J. (2019). The Art of Learning. In T. S. Gendler, & J. Hawthorne (Eds.) *Oxford Studies in Epistemology*, vol. 7. Oxford University Press.

Lewis, D. (1999). Why Conditionalize? In *Papers in Metaphysics and Epistemology*, (pp. 403–407). Cambridge, UK: Cambridge University Press.

Pettigrew, R. (2016). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.

Predd, J., Seiringer, R., Lieb, E. H., Osherson, D., Poor, V., & Kulkarni, S. (2009). Probabilistic Coherence and Proper Scoring Rules. *IEEE Transactions of Information Theory*, *55*(10), 4786–4792.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons.

Schoenfield, M. (2017). Conditionalization does not (in general) Maximize Expected Accuracy. *Mind*, *126*(504), 1155–87.

van Fraassen, B. C. (1989). *Laws and Symmetry*. Oxford: Oxford University Press.

van Fraassen, B. C. (1999). Conditionalization, A New Argument For. *Topoi*, *18*(2), 93–96.

Weisberg, J. (2007). Conditionalization, Reflection, and Self-Knowledge. *Philosophical Studies*, *135*(2), 179–97.