The value of information and the epistemology of inquiry

Richard Pettigrew
Richard.Pettigrew@bristol.ac.uk

June 9, 2024

Contents

Acknowledgements I'm very grateful to Mona Simion, Chris Kelp, Lilith Mace, and Claire Field for organising the Formal meets Normal conference at the Cogito research centre in Glasgow in July 2023; and I'm grateful to the audience there for their excellent questions. This conference prompted me to delve into the literature on inquiry, and the current essay grew out of the talk I gave there.

Version control These notes were originally posted under the title 'Taking a Good look at the norms of gathering and responding to evidence'.

Chapter 1

Introduction

In the recent philosophical literature on inquiry, epistemologists have pointed out that their subdiscipline has typically begun at the point at which you already have your evidence, and it has focussed on identifying the beliefs for which that evidence provides justification or which count as knowledge for someone with that evidence. However, this recent literature goes on to point out, we are not mere passive recipients of the evidence we have.[1] While some of it comes to us unbidden, as we walk along the street, go about our day's work, or chop vegetables for dinner, we often actively collect it. We often choose to put ourselves in positions from which we'll gather some pieces of evidence but not others. We'll move to a position from which we'll see or hear or smell how the world is in one respect but miss how it is in another, as when I choose to inquire into the weather by looking out the window rather than at the forecast on the television screen. We'll prod the world in one way to see how it responds but we won't prod it in another, as when an physicist designs their experiment in this way rather than that. And so on.

As many in the recent literature point out, this has long been recognised, but typically epistemologists have taken the norms that govern inquiry to be practical, not epistemic. We inquire in order to find out things that inform our practical decisions, and so the decision what to find out is governed by practical considerations, and epistemologists leave well alone. Or, even if we inquire in order to find out things without an eye to their practical benefits, the things we do in order to inquire are not the sorts of thing for which one might have epistemic reasons. The recent literature challenges these assumptions and, as a result, uncovers a rich range of questions about the epistemic norms of inquiry.

In this essay, I approach these questions from the so-called formal side of epistemology, and specifically the Bayesian approach that occupies a large part of that side. The starting point for this approach is the observation that

[1] A small sample of recent writings in the epistemology of inquiry: (Hookway, 2006; Friedman, 2020; Kelp, 2021; Thorstad, 2022; Simion, 2023; Flores & Woodard, 2023; Rosa, 2025; Willard-Kyle, forthcoming; Staffel, ta).

beliefs and desires come in different strengths—I believe I'm Scottish more strongly than I believe I'll live to 90; I want this orange more strongly than I want that apple. What's more, we can represent these different strengths numerically, or in some other mathematically tractable way—I'm 99% confident I'm Scottish, but only 10% confident I'll live to 90; the amount by which I prefer an orange to an apple is the same as the amount by which I prefer a blackberry to a gooseberry. When we represent them numerically, we typically call the degrees of belief *credences* and the degrees of desire *utilities*.

Among other things, Bayesians seek norms that govern how you should set your credences initially, how you should change them when new evidence comes in, and how you should combine them with your utilities, and possibly other attitudes, in order to set the preferences that guide your actions. The non-formal side of epistemology, on the other hand, has traditionally been more concerned with full beliefs—ones that don't come in degrees. Among other things, they have been interested in the conditions under which these beliefs have certain features: when they are justified; when they amount to knowledge; when they have the authority to guide action, ground assertion, or figure in further deliberation.

I'd like to persuade you that it's natural to look to the Bayesian approach to help us think about inquiry. For one thing, whether to inquire or not is a choice, as is the decision to inquire in this way rather than that, and Bayesian epistemology offers a number of sophisticated theories of rational choice and decision-making. For another, while inquiring can lead us to adopt full beliefs or abandon them, it can often change our attitudes in more fine-grained ways that leave our full beliefs as they were.[2] Perhaps I am 50% confident in something and my inquiry might make me either 49% confident in it or 51% confident in it, depending on which way the evidence that I collect goes. In this case, I don't have a full belief to begin with, and neither way my evidence goes will lead me to have one afterwards; and yet there is real change, and we must be able to represent that change if we're to understand inquiry, for often the way we come to change our full beliefs is by the gradual accretion of evidence that eventually leads us to believe or disbelieve or suspend judgment.

What's more, because current Bayesian epistemology developed out of the philosophy of science, where it was used to understand how certain sorts of evidence support certain sorts of conclusion, this branch of epistemology has long discussed inquiry. The standard reference in the Bayesian philosophy of science is I. J. Good's (1967) 'On the Principle of Total Evidence' or Leonard J. Savage's (1954) *The Foundations of Statistics*. However, the result at the centre of Good's argument is already proved by Janina Hosiasson in her 1931 *Mind* paper, 'Why do we prefer probabilities relative to many data?'.[3] We know

[2] Arianna Falbo (2023) makes a related point.

[3] I learned of Hosiasson's paper from Christian Torsell's (2024) paper 'Janina Hosiasson and the Value of Evidence', and wholeheartedly recommend those interested in the history of this tradition to consult that. I will refer to her as Janina Hosiasson throughout, since that is the name under which she published this crucial paper. However, she later became Janina Hosiasson-Lindenbaum, and her more influential papers in the then-nascent field of Bayesian philosophy of

from his notes that Frank P. Ramsey (1990) was also aware of it, though he never published it. And, in economics, the standard reference is David Blackwell's (1951) 'Comparison of Experiments', which extends the result considerably in a direction I won't pursue in this essay. All use the same account of the pragmatic utility of gathering evidence, which originates in Hosiasson's paper; this is the basis for all of the future results.

Building on these insights, Bayesian philosophy of science and its more recent expansion into Bayesian epistemology more broadly has produced a reasonably well-developed framework in which to understand norms of inquiry, both epistemic and practical. The first half of this essay presents this framework. In Chapter 2, I will present the framework that has been built on Hosiasson's insight. Its centrepiece is the pragmatic version of the so-called Value of Information Theorem, which formalizes the intuitive idea that you're better off gathering as much information as you can before making a decision, providing the process of gathering it isn't too costly. So, for instance, it's better for me to check the weather forecast before I decide whether or not to take my umbrella with me when I leave the house, providing it doesn't take me too much time. More precisely, it says that, if the information available is free, you'll never do worse in expectation by taking it before making the decision, and you'll very often do better. This tells us something about the pragmatic value of inquiry: it aids decision-making. I'll present the theorem, as well as a series of recent generalizations due to John Geanakoplos (1989), Nilanjan Das (2023), and Kevin Dorst et al. (2021).

I'll also note how things go differently if you use a decision theory other than expected utility theory, such as Lara Buchak's (2013) risk-weighted expected utility or Chris Bottomley and Timothy Luke Williamson's (ta) weighted-linear utility theory; and also how things go if you have imprecise probabilities instead of precise ones.

In Chapter 3, I will present a purely epistemic version of the Value of Information Theorem, due to Graham Oddie (1997). This theorem formalizes the intuitive idea that you'll have better beliefs the more you inquire. So, for instance, my credences about whether it will rain or not will likely be better if I check the weather forecast than if I don't. More precisely, it says that gathering evidence never decreases the accuracy of your credences in expectation, and very often it increases it. This tells us something about the epistemic value of inquiry: it gets us better beliefs. Again, I'll present the theorem, as well as a series of recent generalizations due to Wayne Myrvold (2012), Alejandro Pérez Carballo (2018), and, again, Kevin Dorst et al. (2021).

In the second half of this essay, I put to work the approach to inquiry I've been describing in the first half. I turn to some of the questions from the recent debate about inquiry and ask how the approach initiated by Hosiasson and adapted by Oddie can help us answer them. Questions will include: When should we initiate an inquiry, when should we continue it, when should we conclude it, and when should we reopen it (Chapter 4)? How should we

---

science were published under that name.

understand Julia Staffel's distinction between transitional attitudes and terminal attitudes (Staffel, 2021b) (Chapter 5)? How do epistemic norms of inquiry relate to epistemic norms of belief or credence, and can they conflict, as Jane Friedman (2020) contends? How should we resolve the apparent puzzle raised by Friedman's example of counting the windows in the Chrysler Building? How should we direct our attention, as Georgi Gardiner (2022) asks (Chapter 6)? And how should we understand the epistemic error that occurs when someone is resistant to evidence in the way Mona Simion (2023) describes (Chapter 7)?

Throughout, I will present the ideas both informally and formally. I'll place any formal presentation that uses mathematical notation in blue boxes. These can be skipped over, if you prefer, as the ideas are presented informally in the surrounding text. But I include them, as some people will find it easier to see the ideas presented in that notation, and hopefully some will wish to use this framework and these results themselves, and for that you'll want to be sure that you're formalising things as others have.

Before we begin, I should note that there is another facet to the study of inquiry in formal epistemology: it is the study of collective rather than individual inquiry, and it tends to ask how we should structure our scientific communities, institutions, and practices in order to best discover the truth together.[4] I will not discuss it here, but only because it makes less obvious contact with the questions raised in the recent work on inquiry in non-formal epistemology; though of course it makes important contact with other work in non-formal epistemology, namely, on echo chambers, epistemic bubbles, misinformation, conspiracy theories, and testimony.

[4]See, for instance, (Zollman, 2007, 2010; Rosenstock et al., 2017; Bright, 2017; O'Connor & Weatherall, 2018).

Part I

The value of information

Chapter 2

The pragmatic value of information

While most discussion of the value of evidence or information focuses on what has become known as the Value of Information Theorem, the real contribution lies in the definition of the pragmatic value of gathering evidence, which Ramsey, Hosiasson, Blackwell, and Good all give, however independently. In this chapter, I'll introduce the framework of credences in a little more detail (Section 2.1), give this definition of the pragmatic value of gathering evidence (Section 2.2), and state the Value of Information Theorem (Section 2.3).

2.1 Credences

Hosiasson's definition and the Pragmatic Value of Information Theorem were stated and proved in the standard Bayesian framework, and in this framework we represent someone's beliefs by their *credences*. These are the states we ascribe when we say things like 'Ada is 65% sure it's going to rain' or 'Cal is 50-50 whether they left on the gas'. In these cases, we say Ada has credence 0.65 in the proposition that it's going to rain, and Cal has credence 0.5 in the proposition they left on the gas. Sometimes, these are known as *degrees of belief*, or *strengths of confidence*, *partial beliefs*, or *personal* or *subjective probabilities*. I'll talk of credences throughout.

An individual's credences are numbers assigned to propositions to measure the strength of the individual's beliefs in those propositions. So, what are the propositions? We'll take a pretty flat-footed approach here.[1] We'll assume there's a set of possible worlds that represent the possibilities the individual entertains at the finest level of grain at which they entertain them. And then we'll represent propositions as sets of these possible worlds. And we'll say that the individual assigns credences to propositions thus represented. So, for instance, an individual might only distinguish three states of the world:

- *The Eiffel Tower is taller than the Taj Mahal* (w_1);

[1] Taking this approach is not essential to the framework and results I'm going to present; it'll just make it easier. We could just as well take the objects of credences to be sentences or propositions represented in some other way than as sets of possible worlds.

- *The Eiffel Tower is the same height as the Taj Majal* (w_2);
- *The Taj Mahal is taller than the Eiffel Tower* (w_3).

And then the propositions to which they'll assign credences are those represented by the following sets:

- the empty set ($\emptyset$), which represents the proposition that is true at no world—the necessarily false proposition, if you like;
- the singleton of each of the three states, which represents the proposition true at that state and no other ($\{w_1\}, \{w_2\}, \{w_3\}$);
- the three pairs from the states, which represent *The Eiffel Tower is at least as tall as the Taj Mahal* ($\{w_1, w_2\}$), *The Taj Mahal is at least as tall as the Eiffel Tower* ($\{w_2, w_3\}$), and *The Taj Mahal and the Eiffel Tower are different heights* ($\{w_1, w_3\}$); and
- the set of all three states, which is true at each of them ($\{w_1, w_2, w_3\}$)—the necessarily true proposition.

We call the set of propositions to which the individual assigns credences their *agenda*. For ease of exposition, we assume that there are only finitely many possible worlds they entertain, and so their agenda is similarly finite.

Our individual's *credence function* is then the function that takes each proposition in their agenda and assigns to it the number at least 0 and and at most 1 that represents their credence in it; the number that measures the strength of their belief in it. 0 represents minimal credence or least possible strength of belief; 1 represents maximal credence or greatest possible strength of belief.

We'll assume throughout that our inquiring individual's credence function obeys the Bayesian norm of Probabilism, which says that their credence function must be a probability function. That is, it must assign credence 1 to the necessary truth, namely, the proposition represented by the set of all possibilities; it must assign credence 0 to the necessary falsehood, namely, the proposition represented by the empty set; and the credence it assigns to a disjunction of two mutually exclusive propositions must be the sum of the credences it assigns to the disjuncts.

Suppose W is a finite set of possible worlds, and let $\mathcal{F}$ be the set of subsets of W. Then a *credence function* is a function $C : \mathcal{F} \to [0, 1]$. And a credence function is *probabilistic* iff

(i) $C(\emptyset) = 0$ and $C(W) = 1$;

(ii) $C(X \cup Y) = C(X) + C(Y)$ if $X \cap Y = \emptyset$.

Equivalently,

(i) $\sum_{w \in W} C(w) = 1$;

(ii) $C(X) = \sum_{w \in X} C(w)$.

where we abuse notation and write $C(w)$ for the credence that C assigns to the singleton set $\{w\}$.

Probabilism Rationality requires that your credence function is probabilistic.

2.2 The pragmatic value of gathering evidence

As I mentioned above, the key insight in the Bayesian approach to inquiry is the definition of the pragmatic value of gathering evidence, which was given originally by Janina Hosiasson, Frank Ramsey, David Blackwell, and I. J. Good at different times. Good's is most commonly discussed in philosophy of science, Blackwell's in economics, and Hosiasson's version has been sadly neglected, despite being published first.

This definition begins with another definition; it begins with a definition of the pragmatic value of a probabilistic credence function relative to a decision you will face. Suppose you will face a particular decision between a range of options, where an option is specified by giving its utility at each possible state of the world, and the utility of an option at a world is a real number that measures how much you value the outcome of that option at that world. Then the standard theory of choice under uncertainty says that you should pick an option with maximal expected utility from the point of view of the credence function you have when you face the decision: that is, you calculate the expected utility of each option by taking its utility at each world, weighting that by the credence you assign to that world, and summing up these credence-weighted utilities; and then you should pick an option whose expected utility is maximal—that is, no other option has higher expected utility, though perhaps others have equally high expected utility.

So let's assume you'll do this. Then we can define the pragmatic value for you, at a particular state of the world, of having a particular credence function when faced with a particular decision: it is the utility, at that state of the world, of the option that this credence function will lead you to pick from those available in the decision. This will be one of the options that maximizes expected utility from the point of view of that credence function; and since there might be more than one that maximizes that, we must assume you have a way of breaking ties between them.

The pragmatic utility of a credence function

Some preliminary definitions:[a]

- A *decision problem* $\mathcal{D}$ is a set of options.

- Each *option* o in $\mathcal{D}$ is a function from the set of possible worlds W into the real numbers: $o(w)$ is the utility of o at w.
- Given a probabilistic credence function C, the *expected utility of o from the point of view of C* is $\sum_{w \in W} C(w)o(w)$.
- Given a decision problem $\mathcal{D}$ and a probabilistic credence function C, let $\mathcal{D}_C$ be the *choice set*: that is, it is the set of options in $\mathcal{D}$ that maximize expected utility from the point of view of C.
- A *tie-breaker function* τ takes any set of options and picks one of them out. Our individual uses it when there is more than one option that maximizes expected utility. Given a decision problem $\mathcal{D}$ and a credence function C, they apply τ to the choice set $\mathcal{D}_C$ to give the option $\tau(\mathcal{D}_C)$ that they pick.

Definition 1 (Pragmatic utility of a credence function)**.** *The pragmatic utility, at world w, of a credence function C relative to decision problem $\mathcal{D}$ and tie-breaker function τ, is*

$$\mathrm{PU}^{\mathcal{D},\tau}(C, w) = \tau(\mathcal{D}_C)(w)$$

That is, it is the utility, at w, of $\tau(\mathcal{D}_C)$, which is the option our individual will pick from among those options in $\mathcal{D}$ that maximize expected utility relative to their credence function C.

[a] For our initial presentation of the Value of Information framework, we work within Savage's (1954) version of expected utility theory. In Section 2.4, we consider the status of the Value of Information Theorem in other versions.

So, for instance, suppose I have to walk to the shops and I must decide whether or not to take an umbrella with me. And suppose I have credences concerning whether or not it will rain as I walk there. Let's suppose first that taking the umbrella uniquely maximizes expected value from the point of view of those credences. Then the pragmatic value of those credences at a world at which it does rain is the utility of walking to the shops in the rain with an umbrella, while their pragmatic value at a world at which it doesn't rain is the utility of walking to the shops with no rain carrying an umbrella. And similarly, if leaving without the umbrella uniquely maximizes expected utility from the point of view of those credences, then their pragmatic value at a rainy world is the utility of walking to the shops in the rain without an umbrella, and their pragmatic value at a dry world is the utility of walking to the shops with no rain and no umbrella. And if they both maximize expected utility from the point of view of the credences, then the pragmatic value of the credences will depend on how I break ties. So, holding fixed the decision problem you'll face and the way you break ties, the pragmatic value of a credence function is the utility of the option it'll lead you to pick.

Now, having defined the pragmatic value of a credence function relative to a particular decision you'll face and a way of breaking ties, we can define the pragmatic value of a particular episode of evidence-gathering relative to such a decision and tie-breaker function. We represent such an episode as follows: for each possible state of the world, we specify the strongest proposition you'll learn as evidence at that state of the world—this is what Nilanjan Das (2023) calls an *evidence function*. And we assume that you have a plan for how to respond to each possible piece of evidence—we call this an *updating plan*. Then the pragmatic value, at a particular world, of an episode of evidence-gathering is the pragmatic value, at that world, of the credence function you'll have after learning whatever evidence you'll gather at that world and responding to it in the way your updating plan says you should. So, holding fixed the decision problem you'll face and the way you break ties, the pragmatic value of a credence function is the utility of the option it'll lead you to pick, and the pragmatic value of gathering evidence is the pragmatic value of the credence function it will lead you to have when you respond to that evidence as you plan to.

Of course, this assumes that it is already fixed how you will respond to any evidence you receive; and indeed Good assumes you'll update as the Bayesian says you should: you'll condition on whatever proposition you receive as evidence; that is, your unconditional credence in a proposition after receiving some evidence is your prior conditional credence in that proposition given the evidence you learn, so that my posterior credence in rain after learning the forecast is dry is my prior conditional credence in rain given the forecast says dry. For the moment, we'll stick with this assumption; later, we'll lift it to see what happens.

The pragmatic utility of an evidence-gathering episode

Some preliminary definitions:

- We represent an evidence-gathering episode by an *evidence function* $\mathcal{E} : W \to \mathcal{F}$. This takes each world w to the strongest proposition $\mathcal{E}_w$ you learn in that world if you gather the evidence.
- Given an evidence-gathering episode $\mathcal{E}$ and a prior credence function C, if you engage in the evidence-gathering episode with that prior, your posterior at world w will be $C(- \mid \mathcal{E}_w)$, providing $C(\mathcal{E}_w) > 0$.[a]

Definition 2 (Pragmatic utility of an evidence-gathering episode)**.** *The pragmatic utility, at world w, of an evidence-gathering episode $\mathcal{E}$, relative to decision problem $\mathcal{D}$ and tie-breaker τ, is*

$$\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E}, w) = \mathrm{PU}^{\mathcal{D},\tau}(C(- \mid \mathcal{E}_w), w).$$

That is, it is the utility, at w, of the posterior credence function $C(- \mid \mathcal{E}_w)$ that you will have after learning the evidence you'll learn at that world.

[a] Where $C(X \mid Y) = C(XY)/C(Y)$, when $C(Y) > 0$.

So, for instance, suppose I have to walk to the shops later and, at that point, I'll have to decide whether or not to take an umbrella with me. And suppose that, between now and then, I can gather evidence by looking at the weather forecast. If I do, I'll learn one of two things: rain is forecast, or rain is not forecast. And updating on that evidence as I plan to, should I choose to gather it, might well change my credences concerning whether or not it will rain on my way to the shops. Then what is the value, at a particular state of the world, of gathering evidence by looking at the forecast? Consider a world at which (i) rain is not forecast but (ii) it does rain; and suppose that, upon learning that rain is not forecast, I'll drop my credence in rain low enough that I'll not take my umbrella. Then the value of gathering evidence at that world is the utility of walking to the shops in the rain without an umbrella. In contrast, consider a world at which (i) rain is forecast but (ii) it doesn't rain; and suppose that, upon learning that rain is forecast, I raise my credence in rain high enough that I take the umbrella. Then the value of gathering evidence at that world is the utility of walking to the shops with no rain but carrying an umbrella. And so on.

This is the Hosiasson-Blackwell-Good account of the pragmatic value, at a particular world, of a particular episode of evidence-gathering; and it is relative to the decision problem you'll face with the credences you come to have after updating, and the way you'll break ties between the options, if you need to. With this in hand, we can now define the *expected* pragmatic value of such an episode from the point of view of your prior credence function (or indeed from the point of view of any probability function). And we can also define the expected pragmatic value of not gathering evidence at all, since that is just the degenerate case of evidence-gathering in which you simply learn a tautology at every state of the world.

The expected pragmatic utility of an evidence-gathering episode

The expected utility of an evidence-gathering episode $\mathcal{E}$, from the point of view of C and relative to decision problem $\mathcal{D}$ and tie-breaking function τ, is

$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E})) = \sum_{w \in W} C(w)\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E}, w) = \sum_{w \in W} C(w)\tau(\mathcal{D}_{C(-|\mathcal{E}_w)})(w).$$

The expected utility of not gathering evidence, from the point of view

of C and relative to decision problem $\mathcal{D}$ and tie-breaking function τ, is

$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(C)) = \sum_{w \in W} C(w)\mathrm{PU}^{\mathcal{D},\tau}(C, w) = \sum_{w \in W} C(w)\tau(\mathcal{D}_C)(w).$$

Huttegger and Nielsen's alternative definition

It is worth noting here, before we move on, that Simon Huttegger and Michael Nielsen use a different definition of the expected pragmatic utility of an evidence-gathering episode (Huttegger, 2013, 2014; Huttegger & Nielsen, 2020; Nielsen, 2024). It is equivalent to the one I have given when $\mathcal{E}$ is factive and partitional, but not otherwise. I'll state it here and give an example that motivates using the definition I use instead.

The Huttegger-Nielsen expected utility of an evidence-gathering episode $\mathcal{E}$, from the point of view of C and relative to $\mathcal{D}$ and τ, is

$$\sum_{w \in W} C(w) \sum_{w' \in W} C(w' \mid \mathcal{E}_w)\tau(\mathcal{D}_{C(-|\mathcal{E}_w)})(w')$$

That is, it is the prior's expectation of the posterior's expectations of the choice the posterior will make.

Consider the following case:

Example 1. *I am holding a coloured ball behind my back. You know it is red or yellow or blue, but nothing more than that, and so you divide your credences equally between the three options. You now have the opportunity to learn something. If it's red or yellow, you'll learn it's red or yellow; if it's blue, you'll learn it's yellow or blue. Now consider a bet that pays £10 if the ball is yellow and £0 otherwise and costs £4. Your priors will reject that bet, because the price is too high from their point of view. However, whatever you learn, your posteriors will accept it. The Huttegger-Nielsen definition says that, in this case, you are rationally required to learn the evidence because the option your prior will choose (namely, to reject the bet) has prior expected utility 0, while, from the point of view of whichever posterior you have, the expected utility of the option it will choose (namely, to accept the bet) will be strictly positive, and so the expectation of those expected utilities will be positive. My own view is that, from the point of view of your priors, you should not learn, since learning will lead you to do something for sure that you think is the wrong thing to do. And so I use the definition of expected pragmatic utility of an evidence-gathering episode that I give above, rather than this definition.*

More formally:

- $W = \{Red, Yellow, Blue\}$

- $\mathcal{E}_{Red} = Red \vee Yellow \quad \mathcal{E}_{Yellow} = Red \vee Yellow \quad \mathcal{E}_{Blue} = Yellow \vee Blue$;
- $\mathcal{D} = \{Accept, Reject\}$.

	1/3 *Red*	*1/3* *Yellow*	*1/3* *Blue*
Accept	*-4*	*6*	*-4*
Reject	*0*	*0*	*0*

We are now furnished with all the definitions we need to state Good's Value of Information Theorem.

2.3 The Value of Information Theorem

Good's Value of Information Theorem runs as follows: Fix a decision problem you'll face at a later time; fix the way you'll break ties between a set of options when they all maximize expected utility; and assume you plan to update upon receiving evidence in the way the Bayesian suggests, namely, by conditionalizing on it. Now suppose that, for no cost, you may gather some evidence before you face the decision problem. And suppose that the evidence function that represents this possible evidence-gathering episode is *factive* and *partitional*: to say that it is factive is to say the proposition you'll learn is true; to say it's partitional is to say that the set of propositions you might learn forms a partition—for any possible world, exactly one of these propositions is true at that world. Then the expected pragmatic value, from the point of view of your prior credences, of gathering that evidence is at least as great as the expected pragmatic value, again from the point of view of your prior credences, of not gathering it; and, very often, it is strictly greater. How often? Well, if you assign some positive credence to getting evidence that will result in posterior credences from whose point of view the option that your priors would have chosen is no longer optimal, then the expected pragmatic value of gathering the evidence is strictly greater than the expected pragmatic value of not gathering it. In slogan form: it's always rationally permissible to take free evidence, and it's rationally required when you think it might lead you to consider what you would have chosen without it suboptimal.

Hosiasson's Pragmatic Value of Information Theorem

Some preliminary definitions: Suppose $\mathcal{E}$ is an evidence function. Then:

- $\mathcal{E}$ is *factive* if, for each world w in W, $\mathcal{E}_w$ is true at w;
- $\mathcal{E}$ is *partitional* if $\{E_w : w \in W\}$ forms a partition.

Theorem 1 (The Value of Information Theorem)**.** *If $\mathcal{E}$ is factive and partitional, and C is a probabilistic credence function, then*

(i)
$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E})) \geq \mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(C))$$

(ii)
$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E})) > \mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(C))$$

if there is w in W such that $C(w) > 0$ and $\tau(\mathcal{D}_C) \notin \mathcal{D}_{C(-|\mathcal{E}_w)}$.

In fact, Hosiasson proves a more general result from which the Value of Information theorem follows. To state it, we need a definition:

- *$\mathcal{E}$ is at least as informative as $\mathcal{E}'$* if, for any world w, $\mathcal{E}_w \subseteq \mathcal{E}'_w$.

Lemma 2 (The Value of Information Lemma)**.** *If $\mathcal{E}$ and $\mathcal{E}'$ are factive and partitional, C is a probabilistic credence function, and $\mathcal{E}$ is at least as informative as $\mathcal{E}'$, then*

(i)
$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E})) \geq \mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E}'))$$

(ii)
$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E})) > \mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(\mathcal{E}'))$$

if there is w in W such that $C(w) > 0$ and $\tau(\mathcal{D}_{\mathcal{E}'_w}) \notin \mathcal{D}_{C(-|\mathcal{E}_w)}$.

Theorem 1 follows from Lemma 2 if we note that not gathering evidence is the same as engaging in the degenerate evidence-gathering episode $\mathcal{E}$, where $\mathcal{E}_w$ is the necessary proposition for all w in W, and any factive and partitional evidence function is at least as informative as this degenerate one.

A natural response upon first encountering the Value of Information Theorem is to think it's obviously true. After all, surely any true evidence is guaranteed to improve your epistemic situation, and surely improving the epistemic standpoint from which you face decisions leads to better choices. However, this isn't true. Evidence can be true but misleading. It is easy to find examples in which you would have made a decision that obtained for you more utility had you not learned the true evidence you did before choosing.

Suppose you know it's either sunny, rainy, or windy outside, and you divide your credences equally over the three. In fact, it's windy. You can stay indoors, or you can go outside. Staying indoors gets you 8 units of utility for sure; if you go outside and it's sunny, you get 14 units, if it's windy you get 6, and if it's rainy, you get 1.

	Sunny	Windy	Rainy
Indoors	8	8	8
Outdoors	14	6	1

Then you currently prefer to stay indoors, since the expected utility of doing that (8) is higher than the expected utility of going outside ($\frac{14+6+1}{3} = 7$). Now you learn it's sunny or windy. You update on this information and come to prefer going outside, since your new expected utility for doing that ($\frac{14+6}{2} = 10$) is higher than your new expected utility for staying indoors (which is still 8). But, since it's windy, you end up less well off as a result of learning then choosing; you'd have done better just to choose without learning.

So the Value of Information theorem is not obviously true. So why does it hold? If your evidence will teach you which member of a pre-specified partition is true, while misleading evidence is possible, when its effects are weighted by your credence you'll get it and considered together with the possibility of non-misleading evidence, whose effects are weighted by your credence you'll get them, it turns out that the possibility of non-misleading evidence wins out and it's better, in expectation, to gather the evidence and take the risk. This is helpful to bear in mind as we consider versions of the theorem in other settings in Sections 2.5 and 2.6.

2.4 The Sure Thing Principle

A natural way to see that the Value of Information theorem is true begins with Leonard Savage's *Sure Thing Principle*. This is a putative norm of decision theory—indeed, one of the axioms Savage laid done for rational decision-making. On one reading, it tells us how our preferences after learning something should hang together with our preferences beforehand. Here is the story with which Savage introduces it:

> A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant. So, to clarify the matter to himself, he asks whether he would buy if he knew that the Democratic candidate were going to win, and decides that he would. Similarly, he considers whether he would buy if he knew that the Republican candidate were going to win, and again finds that he would. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say. (Savage, 1954, 21)

The general principle is this: If you would prefer one option (buying the house) to another (not buying it) were you to learn that a given proposition is true (the Democrat wins), and you would prefer that first option to that second option were you to learn that same proposition is false (the Republican wins), then you should prefer the first option to the second before learning whether

it's true or false—here, as before, we assume that, if you learn a proposition, you update your credences by conditioning on it in the way the Bayesian requires you to do.

Standard expected utility theory, which tells you to set your preferences in line with your expected utilities, satisfies this principle, and we can use that fact to show that the Value of Information Theorem is true. Suppose you have the opportunity to learn whether a proposition is true or false before you make a decision. Suppose you learn it's true. Will you be glad you did? That is, will you think the decision you'll now make is at least as good as the decision you would have made had you not learned the proposition? Yes! For you're asking whether your current credences expect the option you'll now choose with those current credences to have at least as great expected utility as the option you would have chosen with your old credences. And of course the answer is that they will. What's more, if your old credences would have led you to choose differently, then your current credences expect the choice you make with them to be strictly better! And the same goes if you were to learn the proposition is false. The credences you'd then have after updating on that information would expect the option you'd choose with them to have maximal expected utility, and be strictly better utility if your old credences would have led you to choose differently. And so, either way, you'd be glad you learned the evidence: and so, by the Sure Thing Principle, prior to learning it, you should prefer learning it to not learning it.

The Sure Thing Principle and Expected Utility Theory

While Savage's example and my discussion of it have stated the Sure Thing Principle informally in terms of what you'd choose now and what you'd choose were you to learn something, the standard statement of principle in fact involves the relationship between your unconditional preference ordering over options and, for each proposition, a conditional preference ordering over options under the supposition of that proposition. This is related to the version concerning learning that I have been discussing informally if we assume that, when we learn a proposition, our new unconditional preferences should be our old conditional preferences under the supposition of that proposition.

We'll write the unconditional preference ordering $\preceq$ and, for each E, the conditional preference ordering $\preceq_E$.

> **The Sure Thing Principle** If $o_1 \preceq_E o_2$ and $o_1 \preceq_{\overline{E}} o_2$, then $o_1 \preceq o_2$.

In expected utility theory, we say that an individual's unconditional preferences order the options by their expected utility relative to their credence function, so that

$$o_1 \preceq o_2 \Leftrightarrow \mathrm{Exp}_C(o_1) \leq \mathrm{Exp}_C(o_2)$$

And we say that their conditional preferences order the options by their expected utility relative to their conditional credence function, so that

$$o_1 \preceq_E o_2 \Leftrightarrow \text{Exp}_{C(-|E)}(o_1) \leq \text{Exp}_{C(-|E)}(o_2)$$

To see that expected utility theory satisfies the Sure Thing Principle, we note the following fact about expected utility theory:

$$\text{Exp}_C(o) = C(E)\text{Exp}_{C(-|E)}(o) + C(\overline{E})\text{Exp}_{C(-|\overline{E})}(o)$$

That is, the expectation of an option is the expectation of its conditional expectations. So if, for both E and $\overline{E}$, the conditional expectation of o_1 is at most the conditional expectation of o_2, then the expectation of the conditional expectations of o_1 must be at most the expectation of the conditional expectations of o_2. And so the expectation of o_1 must be at most the expectation of o_2.

2.5 The value of information in other theories of rational choice

So far, I have been assuming that the correct theory of rational choice is expected utility theory, and indeed I've been using Savage's version of it. But what do alternative decision theories say about the value of gathering evidence?

Suppositional decision theories

I'll begin with the so-called *suppositional decision theories* and describe Michael Nielsen's (2024) elegant result that any such theory that satisfies the value of information must be a particular sort of causal decision theory, namely, one based on what David Lewis (1981) called an *imaging operators*.

In Savage's decision theory, the options between which we choose are represented by functions from worlds to utilities, and we define the value of an option to be its expected utility, that is, the sum over each possible world of the utility of that option at that world weighted by the credence assigned to the world. Notice that, in this theory, the credence assigned to a world does not depend on the option whose value we are assessing. Savage's theory thereby assumes what is sometimes called *act-state independence*: performing the action does not affect the state of the world.

In suppositional theories, we do not assume this. A suppositional theory is defined by a supposition function, which takes the decision-maker's credence function and an option and returns a new credence function, which we think of as giving the decision-maker's credences *under the supposition* that the option is chosen. Then the value of an option according to this theory is just

the expected utility of performing that option under the supposition that it is performed, that is, the sum over each possible world of the utility of that world weighted by the credence in that world under the supposition that the option is chosen. Notice that, in such theories, worlds have the same utility, regardless of which option is chosen. In Savage's theory, in contrast, the utility of the world depends on which option is chosen. In suppositional theories, instead, the option does not alter the utility of the world; it alters the credence in the world that is used to weight the fixed utility of that world.

There are many suppositional theories out there. Richard Jeffrey's (1983) evidential decision theory is a suppositional theory where the credence in a world under the supposition of choosing an option is just the conditional credence in that world given that the option is chosen. Allan Gibbard and William L. Harper's (1978) version of causal decision theory is a suppositional theory where the credence in a given world under the supposition of choosing an option is the probability of the counterfactual that says that if that option were chosen, the world would be that one. And there are others.

Nielsen is particularly interested in the suppositional theories generated by so-called imaging operators. An imaging operator is generated by a *selection function*. A selection function takes an option o and a world w and says, for each possible world w', what proportion of the original credence function's probability in w' should be transferred to w under the supposition that o is performed (a proposition we will write as $\overline{o}$). So, under that supposition, the imaging operator's credence in a world w is the sum, over each possible world w', of the original credence in w', weighted by the selection function's value for the w' under the supposition $\overline{o}$. So the imaging function works through each possible world w', taking a little of its credence to give to w, where the proportion it gives is specified by the selection function.

Nielsen then proves that the Value of Information Theorem holds for a suppositional theory—that is, for every decision problem and every finite partition, the theory values at least as much learning the true element of the partition and then deciding as it does deciding without learning and sometimes more—if, and only if, it is generated by an imaging operator—that is, the supposition function is an imaging operator.

Indeed, Nielsen's result tells us something stronger than that. For he doesn't assume upfront that you learn by conditionalizing on the evidence; rather, he proves that learning by the evidence and choosing in line with a suppositional theory with a supposition function that is an imaging operator is necessary and sufficient for the Value of Information Theorem.

One upshot of Nielsen's result, which we already knew from Brian Skyrms (1990b) is that the Value of Information theorem does not hold of Jeffrey's evidential decision theory, since that cannot be generated by a supposition function generated by an imaging operator.

Suppositional decision theories and Nielsen's theorem

- A *supposition function* s takes a credence function C and a proposition A and returns a credence function $s(C, A)$, where, if $C(w)$ gives the credence in w under no supposition, then $s(C, A)(w)$ gives the credence in world w *under the supposition of* A. We assume $s(C, A)(A) = 1$.
- Given an option o, we write $\overline{o}$ for the proposition that o is performed.
- A decision theory is *suppositional* if it says that you should choose an option with maximal suppositional expected utility relative to a supposition function, where the suppositional expected utility of o relative to supposition function s is given by

$$\sum_{w \in W} s(C, \overline{o})(w) U(w)$$

  where $U(w)$ is the utility of world w.
- Some suppositional theories:

  (i) Evidential decision theory:

$$s(C, A)(w) = C(w \mid A).$$

  (ii) Counterfactual causal decision theory:

$$s(C, A)(w) = C(A \mathrel{\square\!\!\rightarrow} w).$$

  (iii) Stalnakerian causal decision theory:

$$s(C, A)(w) = \sum_{w' \in W} C(w') f(A, w')(w),$$

  where f is an *imaging operator*, that is, $f(A, w')$ is a probability function for each option A and world w', and $f(A, w')(A) = 1$.
- A learning operator l takes a credence function C and a proposition E and returns a credence function $l(C, E)$, where, if $C(w)$ gives the prior credence in w, then $l(C, E)(w)$ gives the posterior credence in world w *upon learning* E. We assume $l(C, E)(E) = 1$.

Theorem 3 ((Nielsen, 2024))**.** *The following are equivalent:*

(i) *For every credence function c, every decision problem $\mathcal{D}$, tie-break function τ, every utility function u, and every factive and partitional evidence function $\mathcal{E}$,*

$$\sum_{E \in \mathcal{E}} C(E) \max_{o \in \mathcal{D}} \sum_{w \in W} s(l(C,E),\overline{o})(w)u(w) \geq \max_{o \in \mathcal{D}} \sum_{w \in W} s(C,\overline{o})(w)u(w)$$

where we abuse notation and write $\mathcal{E}$ to be the set of propositions you might learn—i.e. $\{E$: there is w such that $E = \mathcal{E}_w\}$.

(ii) (a) $l(C,E)(w) = C(w \mid E)$ *whenever* $C(E) > 0$*; and*

(b) $s(C,A)(w) = \sum_{w' \in W} C(w')f(A,w')(w)$*, for some imaging operator f.*

It should be noted that Nielsen is here using the account of the expected pragmatic value of an evidence-gathering episode that he and Simon Huttegger favour and that I raised concerns about in Example 1. That is, Nielsen takes the value of an evidence-gathering episode to be the prior's expectation of its posterior value—that is, the prior's expectation of the posterior's expectation of the act the posterior favours. The account I favour takes the value to be the prior's expectation of the plan to choose whatever your posterior favours.

However, it turns out that, just as the two accounts are equivalent for factive and partitional evidence in Savage's framework, so they are in this framework, providing we assume something about the suppositional expected value of a plan. Here, I'm taking a plan to be a function that takes a world and returns the act you'll pick at that world. The plan that interests us here is the plan to pick whatever action maximizes suppositional expected utility relative to the posterior you'll have at that world once you receive the evidence and update. Let's call that plan *Learn*. What's its suppositional expected utility? Here's the assumption we need:

$$s(C, Learn) = \sum_w C(w)s(C(- \mid \mathcal{E}_w), \tau(\mathcal{D}_{l(C,\mathcal{E}_w)})$$

Risk-sensitive and ambiguity-sensitive decision theories

Shortly after Savage formulated the Sure Thing Principle, two challenges to it were raised: the so-called Ellsberg paradox and the so-called Allais paradox (Allais, 1953; Ellsberg, 1961). In each case we have a pair of decision problems, and when faced with those problems, people tend to choose in ways that are incompatible with the Sure Thing Principle. And yet we are hesitant to say

that those choices are irrational. So there seem to be rational preferences that violate the Sure Thing Principle. And because they violate that, the Value of Information Theorem does not hold for them.

The Ellsberg paradox (Ellsberg, 1961). Before you sits an urn filled with red, black, and yellow balls. There are 90 balls in total. Exactly 30 are red. The remaining 60 are black or yellow, but you do not know in what proportion. A ball is about to be drawn from the urn. You are offered two different choices: the first between Option A and Option B; the second between Option C and Option D. The outcomes of these options depend on the colour of the ball. They're given in the following table:

	Yellow	Black	Red
Option A	£0m	£0m	£1m
Option B	£0m	£1m	£0m
Option C	£1m	£0m	£1m
Option D	£1m	£1m	£0m

Experimentally, Ellsberg found that people strictly prefer A to B, and D to C. The most common explanation is that they are averse to ambiguity. People prefer gambles for which they know the objective expected values. So they prefer A to B because they know the objective probability of Red (one-third) and the objective probability of Yellow or Black (two-thirds); and that is enough to calculate the objective expected value of A, but it does not fix the objective expected value of B, since the objective probability of Black could be anywhere from zero to two-thirds. Similarly, they prefer D to C because they know the objective probability of Red (one-third) and the objective probability of Yellow or Black (two-thirds); and that is enough to calculate the objective expected value of D, but not to calculate the objective expected value of C, since the objective probability of Yellow or Red could be anywhere from one-third to one, while the objective probability of Black could be anything from zero to two-thirds.

These preferences—often referred to as the Ellsberg preferences—violate the Sure Thing Principle. To see this, we reason by reductio. So suppose you have the Ellsberg preferences and you do satisfy the Sure Thing Principle. If you learn that the ball drawn is yellow, you'll be indifferent between A and B since they have the same outcome in that situation. So that means that, if you learn the ball drawn is not yellow, the Sure Thing Principle demands you must strictly prefer A to B, since you prefer A to B unconditionally. But now look at the second pair of options. If you learn that the ball drawn is Yellow, you'll be indifferent between C and D since they have the same outcome in that situation. So that means that, if you learn the ball drawn is not yellow, the Sure Thing Principle demands you must strictly prefer D to C, since you prefer D to C overall. But, if you learn the ball isn't yellow, then A is guaranteed to have the same outcome as C and B is guaranteed to have the same outcome as D. And so, if you learn the ball isn't Yellow, you should prefer A to B iff you

prefer C to D. But that isn't true. We have a contradiction. And so you don't satisfy the Sure Thing Principle.

And we can translate this violation of the Sure Thing Principle into a failure of the Value of Information. Suppose (i) you prefer A to B, (ii) if you learn the ball is yellow, you'll be indifferent between A and B, and (iii) if you learn the ball isn't yellow, you'll strictly prefer B to A. Then you'll pay to avoid learning whether or not the ball is yellow. And similarly, suppose (i) you prefer D to C, (ii) if you learn the ball is yellow, you'll be indifferent between C and D, and (iii) if you learn the ball isn't yellow, you'll strictly prefer C to D. Then again you'll pay to avoid learning whether or not the ball is yellow. But, as we saw above, one of these must be true.

One upshot of this is that decision theories designed to accommodate the Ellsberg preferences will not always value gathering evidence, even when the evidence function is factive and partitional. One such decision theory is Γ-Maximin (Berger, 1985). Suppose your evidence rules out some objective probability functions, but not all. Then Γ-Maximin says you should pick an option whose minimal possible objective expected value among the objective probability functions compatible with your evidence is maximal.

The Allais paradox (Allais, 1953). You hold a ticket for a lottery in which there are 100 tickets. Before you learn which number is on our ticket, you are offered two different choices: the first between Option A′ and and Option B′; the second between Option C′ and Option D′. The outcomes of these options depend on the number on your ticket. They're given in the following table:

	Ticket 1-89	Ticket 90	Ticket 91-100
Option A′	£1m	£1m	£1m
Option B′	£1m	£0m	£5m
Option C′	£0m	£1m	£1m
Option D′	£0m	£0m	£5m

Allais claimed that many people prefer A′ to B′ and D′ to C′, and moreover that this pair of preferences is rational. This idea is that, faced with the first choice, a risk-averse person will prefer A′ over B′ because, while B′ gives a possibility of even greater wealth, the possibility of getting nothing makes B′ less desirable than A′ overall—A′ at least gains you something for sure. On the other hand, when faced with the choice between C′ and D′, where there is no option that guarantees a gain, the greater gain made possible by D′ seems worth it.

Again, and by exactly analogous reasoning, we can see that these preferences—the so-called Allais preferences—violate the Sure Thing Principle. And from that we can infer that, in either the choice between A′ and B′ or in the choice between C′ and D′, they do not value gathering the evidence whether the ticket number is between 1 and 89 or between 90 and 100.

One upshot of this is that decision theories designed to accommodate the Allais preferences will not always value gathering evidence, even when the evidence function is factive and partitional. These include John Quiggin's

(1982; 1993) *rank-dependent utility theory*, Lara Buchak's (2013) *risk-weighted expected utility theory*, and Chris Bottomley and Timothy Luke Williamson's (ta) *weighted-linear utility theory*.

2.6 Generalizing the pragmatic Value of Information theorem

Now, the Value of Information Theorem as I have introduced it is severely limited in application: (1) evidence is rarely free; (2) inquiry involves not only deciding whether or not to gather a specific sort of evidence, but whether to gather this piece of evidence or that piece or to do something entirely different; (3) we rarely know exactly which decision we will face using our credences after the evidence is gathered; (4) evidence doesn't always tell you which member of a pre-specified partition is true; and (5) we'd like some reassurance that, when we do learn whatever we learn, the Bayesian command to update by conditioning on the evidence is the right one. In this section, we address these shortcomings.

#1: Factoring in the cost of evidence. While the Value of Information Theorem is interesting, the real value of the framework that Hosiasson, Blackwell, and Good introduced lies in the account of the pragmatic utility of an evidence-gathering episode at a world. And so, if there is a cost to gathering a certain sort of evidence, we can simply subtract the utility of whatever it is that it will cost us from the utility that gathering it gains for us to give the true pragmatic utility of gathering a specific piece of evidence; this even works if the cost is different in different worlds. And then we can take the expectation of this true pragmatic utility that factors in the cost, and compare it to the pragmatic utility of not gathering the evidence, which we can usually assume is cost-free.

#2: Comparing different evidence-gathering episodes. This account of the true pragmatic utility of gathering some evidence allows us to compare the expected utility of gathering *that* evidence with *that* cost to the expected utility of gathering *this alternative* evidence with *this* cost. After all, in inquiry, our choices are rarely simply to gather some evidence or not; they are choices between different evidence we might gather as well as other sorts of options; and the different sorts of evidence we might gather might have different costs. So, for instance, I might go to the window to see how the sky looks to inform my decision whether or not to take an umbrella, or I might look at the weather app on my smartphone, or I might do both, and each of these options might have different attendant costs.

In fact, Hosiasson's original paper, and Good's later one, both address a version of this question. Suppose you might gather evidence in one way, and it will teach you what the temperature is in Baltimore; or you might gather evidence in another way, and it will teach you something strictly more informative, such as what the temperature is in Baltimore *as well as* what the

temperature is in Boston. Then, if the two evidence-gathering episodes have the same cost, the expected utility of engaging in the more informative one is at least the expected utility of engaging in the less informative one; what's more, that expected utility will be greater if you assign a positive credence to a world in which acquiring the less informative evidence would lead you to choose an option that is suboptimal from the point of view of the credences you would obtain at that world if you were to gather the evidence at that world.

As well as cases in which we wish to compare different evidence-gathering episodes we might undertake, we might also wish to compare an evidence-gathering episode and an alternative option that isn't an evidence-gathering episode at all. And the account of pragmatic value of evidence-gathering that we obtain from Hosiasson allows us to compare them as well. Perhaps I could check the sky from the window, check the weather app, do both, or I could do something else completely, such as making a sandwich for lunch. I can compare the expected utility of each and choose on that basis.

This allows us to consider the so-called opportunity cost of gathering a particular piece of evidence. This is not a cost that we factor into the pragmatic utility of each evidence-gathering episode. Rather, the opportunity cost incurred by doing one thing is the utility we would have got if we'd done some other thing instead. So the opportunity cost of gathering some evidence when I could have made a sandwich for lunch is whatever utility I would have got from making that sandwich. And the opportunity cost of gathering this evidence rather than that is the utility I would have got if I'd gathered that evidence instead.

#3: Allowing uncertainty about the decision problem you'll face. To define the pragmatic utility of an evidence-gathering episode, Hosiasson assumes you know for sure which decision you'll face using your credences, but of course you might be uncertain about this. Fortunately, it's easy to incorporate this uncertainty and recover a version of the Value of Information Theorem. We begin by ensuring that our possible worlds specify not only the truth values of the propositions to which we assign credences, but also which decision we'll face with our credences. We then assign credences to these more fine-grained possible worlds. And, having done all this, we define the pragmatic value of a credence function at a fine-grained world to be the utility at that world of the option it would lead us to choose from the decision we face at that world *once it updates on what decision problem we face at that world*. This latter clause is crucial: if we omit it, and fail to update on the decision problem we face when we become clear about which it is, then the Value of Information Theorem fails. So, the pragmatic value of an evidence-gathering episode is the pragmatic value of the credence function you'll end up with after gathering the evidence, updating on it, and learning what decision problem you face, and updating on that. With this amendment, the Value of Information Theorem still goes through.

Uncertainty about the decision problem

- Suppose **D** is a finite set of possible decision problems you might face.
- Suppose your credence function C is defined over the full algebra of subsets of $W \times \mathbf{D} = \{(w, \mathcal{D}) : w \in W \,\&\, \mathcal{D} \in \mathbf{D}\}$.
- Then define

$$\begin{aligned} \mathrm{Exp}_C(\mathrm{PU}^\tau(\mathcal{E})) &= \sum_{\substack{w \in W \\ \mathcal{D} \in \mathbf{D}}} C(w \,\&\, \mathcal{D}) \tau(\mathcal{D}_{C(-|\mathcal{E}_w \,\&\, \mathcal{D})})(w) \\ &= \sum_{\mathcal{D} \in \mathbf{D}} C(\mathcal{D}) \sum_{w \in W} C(w \mid \mathcal{D}) \tau(\mathcal{D}_{C(-|\mathcal{E}_w \,\&\, \mathcal{D})})(w) \end{aligned}$$

and

$$\begin{aligned} \mathrm{Exp}_C(\mathrm{PU}^\tau(C)) &= \sum_{\substack{w \in W \\ \mathcal{D} \in \mathbf{D}}} C(w \,\&\, \mathcal{D}) \tau(\mathcal{D}_{C(-|\mathcal{D})})(w) \\ &= \sum_{\mathcal{D} \in \mathbf{D}} C(\mathcal{D}) \sum_{w \in W} C(w \mid \mathcal{D}) \tau(\mathcal{D}_{C(-|\mathcal{D})})(w) \end{aligned}$$

Then

Theorem 4. *If $\mathcal{E}$ is factive and partitional, and C is a probabilistic credence function, then*

(i)
$$\mathrm{Exp}_C(\mathrm{PU}^\tau(\mathcal{E})) \geq \mathrm{Exp}_C(\mathrm{PU}^\tau(C))$$

(ii)
$$\mathrm{Exp}_C(\mathrm{PU}^\tau(\mathcal{E})) > \mathrm{Exp}_C(\mathrm{PU}^\tau(C))$$

*if there is w in W and $\mathcal{D}$ in **D** such that $C(w \,\&\, \mathcal{D}) > 0$ and $\tau(\mathcal{D}_{C(-|\mathcal{D})}) \notin \mathcal{D}_{C(-|\mathcal{E}_w \,\&\, \mathcal{D})}$.*

The following example shows why we must insist that the individual update on the decision problem they face before choosing.

Example 2. *Suppose:*

- $W = \{w_1, w_2, w_3\}$;
- $\mathcal{E}_{w_1} = \{w_1, w_2\} \quad \mathcal{E}_{w_2} = \{w_1, w_2\} \quad \mathcal{E}_{w_3} = \{w_3\}$;
- $\mathcal{D} = \{o_1, o_2\} \quad \mathcal{D}' = \{o_1', o_2'\}$.

	w_1	w_2	w_3
$C(-)$	1/3	1/3	1/3
$C(- \mid \mathcal{E}_{w_1})$	1/2	1/2	0
$C(- \mid \mathcal{E}_{w_2})$	1/2	1/2	0
$C(- \mid \mathcal{E}_{w_3})$	0	0	1
$C(- \mid \mathcal{D})$	1	0	0
$C(- \mid \mathcal{D}')$	1/4	3/8	3/8
$o_1(-)$	-2	3	-2
$o_2(-)$	0	0	0
$o_1'(-)$	1	1	1
$o_2'(-)$	0	0	0

Then:

- *$C(- \mid \mathcal{E}_{w_1})$ will pick o_1 when faced with $\mathcal{D}$*
- *C will pick o_2 when faced with $\mathcal{D}$;*
- *C is certain of w_1 conditional on $\mathcal{D}$;*
- *so, C strictly prefers not learning to learning conditional on $\mathcal{D}$.*

What's more:

- *C, $C(- \mid \mathcal{E}_{w_1})$, $C(- \mid \mathcal{E}_{w_2})$, and $C(- \mid \mathcal{E}_{w_3})$ will all pick o_1' when faced with $\mathcal{D}'$;*
- *so, C is indifferent between learning and not learning conditional on $\mathcal{D}'$.*

Therefore, since C assigns some positive credence to $\mathcal{D}$,

- *C strictly prefers not learning.*

#4: Allowing non-factive, non-partitional evidence. As stated, the Value of Information Theorem only covers cases in which the evidence-gathering episode will teach you which element of a partition is true. This is very idealized, though it is faithful to a certain way in which we gather evidence in science. When I measure the weight of a chemical sample, or when I ask how many organisms in a given population are infected after exposure to a particular pathogen, there is a fixed partition from which my evidence will come: I'll learn the sample is this weight or that weight or another one; I'll learn the number of infected organisms was zero or one or two or...up to the size of the population. But of course there are many cases in which our evidence-gathering will not be partitional or even factive in this way. Does the Value of Information theorem hold for any evidence function? If not, can we find weaker conditions on evidence-gathering episodes such that the theorem still holds?

First, we can easily see that there are evidence functions for which the Value of Information theorem does not hold. Take, for instance, a case in which you'll learn the opposite of what is true. If it's going to rain, you'll learn it won't, and if it won't, you'll learn it will. Then it's not hard to see that lots of priors will consider it worse in expectation to gather this evidence than not to, and for lots of decision problems. But there are less extreme cases, and in particular there are ones in which the evidence you receive is factive, unlike in the case just given. Here's one: You know the handkerchief in my pocket is rose, teal, or ochre, but you've no further information, so you assign the same credence to each. You can ask me what colour it is, and if it's rose, I'll say it's rose or teal, if it's teal, I'll say it's rose or teal, and if it's ochre, I'll say it's teal or ochre. And, later you'll face the following decision: you can choose a gamble, which will gain you three units of utility if it's teal and lose two if it's rose or ochre; or you can choose the sure thing, which will win you nothing and lose you nothing whatever colour it is. Then your current credences expect themselves to choose better in this situation than they expect your future credences to choose should you gather the information and update on it. To see this, note that your current credences will take the sure thing, whereas whatever you learn from the evidence-gathering, your updated credences will take the gamble.

Example 3 (The Colour of the Handkerchief)**.** *Suppose:*

- $W = \{w_1 = \textit{Rose}, w_2 = \textit{Teal}, w_3 = \textit{Ochre}\}$;
- $\mathcal{E}_{w_1} = \{w_1, w_2\} = \textit{Rose} \vee \textit{Teal}$;
- $\mathcal{E}_{w_2} = \{w_1, w_2\} = \textit{Rose} \vee \textit{Teal}$;
- $\mathcal{E}_{w_3} = \{w_2, w_3\} = \textit{Teal} \vee \textit{Ochre}$;
- $\mathcal{D} = \{o_1, o_2\}$.

	w_1	w_2	w_3
$C(-)$	1/3	1/3	1/3
$C(- \mid \mathcal{E}_{w_1})$	1/2	1/2	0
$C(- \mid \mathcal{E}_{w_2})$	1/2	1/2	0
$C(- \mid \mathcal{E}_{w_3})$	0	1/2	1/2
$o_1(-)$	−2	3	−2
$o_2(-)$	0	0	0

Then, C will pick o_2, while $C(- \mid \mathcal{E}_{w_1})$, $C(- \mid \mathcal{E}_{w_2})$, and $C(- \mid \mathcal{E}_{w_3})$ will all pick o_1. So, C prefers not to learn the evidence.

So the Value of Information Theorem does not hold for all possible evidence-gathering scenarios, and even fails for reasonably quotidian ones—we'll see more examples below. So the question arises whether we can weaken the conditions of factivity and partitionality to understand better when it does hold.

I'll describe three such attempts:

John Geanakoplos (1989) gives conditions on the evidence-gathering episode itself, and shows that, if it satisfies those, then for any prior credence function you have and any decision problem you'll face, providing you plan to update your prior by conditioning on whatever evidence you learn, gathering the evidence is never worse and often better than not gathering, in expectation. Nilanjan Das (2023) does something similar.

Geanakoplos' strengthening of the Value of Information theorem

Some preliminary definitions: Suppose $\mathcal{E}$ is an evidence function. Then:

- $\mathcal{E}$ is *factive* if $w \in \mathcal{E}_w$, for all w in W.

  That is, whatever evidence you receive will be true.

  The evidence in Example B below isn't factive.

- $\mathcal{E}$ is *positively introspectible* if, whenever $w_2 \in \mathcal{E}_{w_1}$ and $w_3 \in \mathcal{E}_{w_2}$, then $w_3 \in \mathcal{E}_{w_1}$.

  That is, if your evidence at one world leaves another world open, and your evidence at the second world leaves a third world open, then your evidence at the first world should leave the third world open. But we can paraphrase it more intuitively as follows: if a certain proposition gives the strongest proposition you learn, then you also learn that the strongest proposition you've learned entails this proposition. After all, if $\mathcal{E}$ is positively introspectible, then, whatever world you're at, your evidence rules out all worlds at which the evidence you'd have there doesn't entail the evidence you actually have. So, if your actual evidence is E, then at all worlds at which E is true, your evidence at those worlds entails E.

  The evidence in Example C below isn't positively introspectible.

- $\mathcal{E}$ is *nested* if for any w_1 and w_2, either (i) $\mathcal{E}_{w_1} \subseteq \mathcal{E}_{w_2}$, (ii) $\mathcal{E}_{w_2} \subseteq \mathcal{E}_{w_1}$, or (iii) $\mathcal{E}_{w_1} \cap \mathcal{E}_{w_2} = \emptyset$.

  That is, if your evidence at two worlds overlaps, then one must entail the other.

  The evidence in Example C below isn't nested; nor is the evidence in the example of the Colour of the Handkerchief.

Example A below is factive, positively introspectible, and nested, but it is not partitional.

Theorem 5 ((Geanakoplos, 1989)). *If $\mathcal{E}$ is factive, positively introspectible, and nested, then for any prior credence function C, decision problem $\mathcal{D}$, and tie-breaking function τ,*

$$\text{Exp}_C(\text{PU}^{\mathcal{D},\tau}(\mathcal{E})) \geq \text{Exp}_C(\text{PU}^{\mathcal{D},\tau}(C))$$

with a strict inequality if there is w in W such that $C(w) > 0$ and $\tau(\mathcal{D}_C) \notin \mathcal{D}_{C(-|\mathcal{E}_w)}$.

Recall the example from above of the coloured handkerchief—I formalized it in Example 3. A key feature of that example is that the prior credences don't satisfy the Weak Reflection Principle with respect to the possible posteriors you might adopt after gathering the evidence.[2] This says that your prior credence function should be a weighted average of those possible evidence-informed posterior credence functions: that is, there should be some set of weights, one for each possible posterior, such that your credence in a given proposition is the weighted sum of the credences in that proposition assigned by the possible posteriors. You can tell that the credences in the handkerchief example violate this principle because the posterior credence in the handkerchief being teal is $1/2$ whatever you learn in the evidence-gathering episode, while the prior credence is $1/3$—no weighted sum of multiple $1/2$s gives $1/3$! The crucial fact is this: if your prior isn't a weighted sum of your possible posteriors, then there is a choice between two options in which your prior will choose one way, while all of your possible posteriors will choose the other way.[3] And in this case, it's clear that your prior will prefer not to gather evidence and stick with its own judgments rather than gather the evidence and thereby take on one of the posteriors to which that leads.

Weak Reflection Principle and the pragmatic value of information

A preliminary definition:

- An *updating plan* is a function R that takes a possible world w and returns a credence function R_w. The idea is that R_w is the credence function that R endorses at world w.

  If C is your prior, $\mathcal{E}$ is your evidence function, and you plan to update by conditionalization, then your updating plan will be $R_w(-) = C(- \mid \mathcal{E}_w)$. But clearly there are other possible updating plans.

[2] The Weak Reflection Principle is my name for it. The principle itself is due to Bas van Fraassen (1999), who also formulated the Strong Reflection Principle we'll meet below.

[3] I state this result as Theorem 1 in (Pettigrew, 2023a), but it is not original to me.

- The pragmatic utility of an updating plan R at a world w is

$$\mathrm{PU}^{\mathcal{D},\tau}(R, w) = \mathrm{PU}^{\mathcal{D},\tau}(R_w, w).$$

Weak Reflection Principle Suppose C is your prior credence function and R is your updating plan. Then there should be weights $\lambda_w \geq 0$, one for each w in W, such that $\sum_w \lambda_w = 1$ and

$$C(-) = \sum_w \lambda_w R_w(-)$$

Theorem 6. *If C, R do not satisfy the Weak Reflection Principle, then there is a decision problem $\mathcal{D}$ such that, for any tie-breaking function τ,*

$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(R)) < \mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(C))$$

In particular, $\mathcal{D} = \{o_1, o_2\}$ and

(i) *C prefers o_1 to o_2,*

(ii) *R_w prefers o_2 to o_1, for all w in W.*

So, if you violate the Weak Reflection Principle, there is a decision problem you might face such that your prior will prefer to make the decision itself rather than gather evidence and hand over to the resulting posteriors. But notice: satisfying the Weak Reflection Principle is certainly not sufficient to avoid this. Think of someone who is 50-50 about whether it will rain or not; if they choose to gather the evidence and it's raining, they'll learn it's not, but if they choose to gather it and it's not, they'll learn it is—so they're evidence is perfectly anti-factive. Their priors nonetheless satisfy the Weak Reflection Principle. So we need something stronger if we want a condition sufficient to make gathering evidence preferable to not gathering it.

One stronger condition that is sufficient is van Fraassen's original Reflection Principle, which I'll call the Strong Reflection Principle here (van Fraassen, 1984, 1995). This says that your prior credence function, conditional on your posterior being a particular probability function, is just that probability function. If your prior satisfies that principle with respect to your possible posteriors, then you will prefer to gather the evidence that gives your posteriors.

Strong Reflection Principle and pragmatic value of information

A preliminary definition:

- If R is your updating plan and P is in $\{R_w : w \in W\}$, then write $\overline{P}$ for the proposition that is true at all worlds w in W such that

$P = R_w$. That is, $\overline{P}$ says that P is the credence function your updating plan endorses at your world.

Strong Reflection Principle Suppose C is your prior credence function and R is your updating plan. Then for all P in $\{R_w : w \in W\}$, the following should hold:

$$C(- \mid \overline{P}) = P(-)$$

if $C(\overline{P}) > 0$.

Theorem 7. *If C, R satisfy the Strong Reflection Principle, then, for any decision problem $\mathcal{D}$ and tie-breaking function τ,*

$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(R)) \geq \mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(C))$$

with a strict inequality if there is w in W such that $C(w) > 0$ and $\tau(\mathcal{D}_C) \notin \mathcal{D}_{R_w}$.

However, as Kevin Dorst et al. (2021) note, there are credence functions that satisfy the Weak Reflection Principle but not the Strong Reflection Principle and yet for which the Value of Information Theorem holds. And we know there are credence functions that satisfy the Weak Reflection Principle for which the theorem does not hold. And so they seek a principle that lies somewhere between the two: stronger than the Weak version and weaker than the Strong one.

Enter the principle of Total Trust. Suppose we have what statisticians call a *random variable*. This is a quantity that can take different values at different possible worlds—it might be the amount of rainfall in Santiago next May, for instance. Now, suppose we take some threshold value—let's say 5mm—and consider the proposition that says that the expectation of this quantity from the point of view of your posterior credence function takes a value no less than this threshold—that is, it says that your posterior credence function expects the rainfall in Santiago next May to be at least 5mm. Now condition your prior on that proposition and, from the point of view of the resulting probability function, calculate the expected value of that quantity—so now we're talking about the expectation of the rainfall in Santiago next May from the point of view of your prior credence function once it's been conditioned on the proposition that your posterior expects that rainfall to be at least 5mm. Then Total Trust says that this expectation should also take a value no less than the threshold—your prior credence function conditional on the claim that your posterior will expect the rainfall to be no less than 5mm should expect the rainfall to be no less than 5mm. And this should hold for any quantity whatsoever—not just Santiago's rainfall. If your prior satisfies Total Trust, then the Value of Information Theorem holds.

Total Trust Principle and pragmatic value of information

Some preliminary definitions:

- Given an updating plan R, a random variable X, and a real number t, let $\langle \mathrm{Exp}_R(X) \geq t \rangle$ be the proposition that is true at all worlds w for which $\mathrm{Exp}_{R_w}(X) = \sum_{w' \in W} R_w(w')X(w') \geq t$.

Total Trust Principle Suppose C is your prior credence function and R is your updating plan. Then, for any random variable X and any threshold t, the following should hold:

$$\mathrm{Exp}_C(X \mid \mathrm{Exp}_R(X) \geq t) \geq t.$$

Theorem 8 ((Dorst et al., 2021))**.** *The following are equivalent:*

(i) *C, R satisfy the Total Trust Principle*

(ii) *For any decision problem $\mathcal{D}$ and tie-breaking function τ,*

$$\mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(R)) \geq \mathrm{Exp}_C(\mathrm{PU}^{\mathcal{D},\tau}(C))$$

with a strict inequality if there is w in W such that $C(w) > 0$ and $\tau(\mathcal{D}_C) \notin \mathcal{D}_{R_w}$.

I won't delve any further into the details of the general results here; instead, I'll note a couple of examples that illustrate how many ways factivity and partitionality can fail while retaining the value of evidence-gathering.

A. Good and bad cases. In discussions of scepticism, whether it concerns the external world, other minds, or something else, externalists often distinguish themselves from internalists by claiming that the evidence you'd have in the 'good' or non-sceptical situation is different from the evidence you'd have in the 'bad' or sceptical situation (Williamson, 2013). In the good situation, your evidence is that you're in the good situation, while in the bad situation, your evidence is that you're either in the good situation or in the bad situation. Suppose I can gather evidence of this sort, perhaps by meeting another person about whom I am currently uncertain whether they have a mind, and then make a decision afterwards. Should I? The evidence in this case is factive, but non-partitional, since the evidence in the bad situation overlaps with the evidence in the good situation. Nonetheless, it satisfies the weaker conditions that Geanakoplos (1989) enumerates, and so it is always better in expectation to gather this evidence if you think it might lead you to change your mind about what to choose. Figure 2.1 illustrates the evidence function in this case.

B. Misdirection vs complete information. Someone in a company has committed fraud and it's your job to find out who it is. There are three suspects: the CEO, the COO, and the CFO. You have the opportunity to interview the

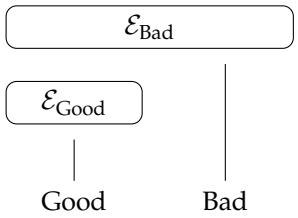


Figure 2.1: The evidence function in good/bad cases (Example A).

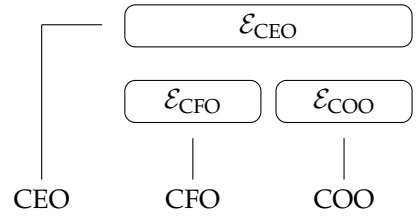


Figure 2.2: The evidence function in the CEO/COO/CFO case (Example B).

CEO's assistant, and you know they know who did it. But you also know they're deeply loyal to the CEO. So, if it's the CFO, they'll tell you that; if it's the COO, they'll tell you that; but if it's the CEO, they'll tell you it's the CFO or the COO. So in this case, the evidence is not factive and it's not partitional. Figure 2.2 illustrates the evidence function in this case.

This is the sort of case that Nilanjan Das (2023) and Bernhard Salow (2018) call a *biased inquiry*, since there is a proposition such that you know your credence in it will rise regardless of what you learn, namely, the proposition that it was the CFO or the COO. This is a particular way in which you might violate the Weak Reflection Principle from above. And so we know that, for any prior that gives positive credence to each of the possibilities—CEO, COO, CFO—there is a decision problem your priors will prefer to face using themselves rather than the posteriors they'd get from gathering the evidence.

But that's not necessarily true of all decision problems. If you have a reasonably high prior that it's the CEO or if there's a big difference in the utilities of the different options at the state of the world at which it is the CEO, then you should not take the evidence, since it's too misleading relative to your prior and the decision problem. But if there's no difference between the utility of the options at the world at which the CEO is guilty, perhaps because you know there's no way to prosecute that individual anyway, then you should take the evidence, since it gives the opportunity of learning exactly who did it if it's the one of the other two. So this is a very clear case in which you have to weigh up misleading evidence, which you'll receive if the CEO is guilty, against highly accurate and informative evidence, which you'll receive if the CEO is innocent. How you weigh it up depends on your priors, but also the decision you face.

Example 4 (Fraud in the company)**.**

- $W = \{w_1 = CEO, w_2 = CFO, w_3 = COO\}$;
- $\mathcal{E}_{w_1} = \{w_2, w_3\} = CFO \vee COO$;
- $\mathcal{E}_{w_2} = \{w_2\} = CFO$;
- $\mathcal{E}_{w_3} = \{w_3\} = COO$.

$C(-)$	w_1 1/3	w_2 1/3	w_3 1/3
o_1	1	4	2
o_2	1	2	3
o_1'	−3	1	1
o_2'	0	0	0

Then:

- *C prefers to learn the evidence before facing* $\mathcal{D} = \{o_1, o_2\}$.

  After all, there is no difference between o_1 *and* o_2 *in the world in which they'll get false evidence, so it doesn't matter what they'll pick; and the evidence will lead them to pick the best for sure in the other worlds.*

- *C prefers not to learn the evidence before facing* $\mathcal{D}' = \{o_1', o_2'\}$.

  After all, C prefers o_2' *to* o_1'*, while each possible posterior prefers* o_1' *to* o_2'.

At this point, it might occur to you to ask: is there no limit to the falseness of evidence we might sometimes prefer to acquire? In the case just described, you weighed up the possibility of false evidence against the possibility of very informative true evidence. But could there be a case in which all the possible evidence is false and yet you'd still choose to gather it? The answer is yes. Suppose there are four states of the world, and you must choose between two options. The first gives zero units of utility for sure, while the second gives positive utility at two worlds and negative utility at two worlds. Then your ideal situation would be to have credences that choose the first option at the worlds at which the second has negative utility and the second option at worlds at which the second has positive utility. Now suppose that, if you're in one of the worlds at which the second option as negative utility, you'll learn you're at the other world at which it has negative utility; and if you're in one of the worlds at which the second option has positive utility, you'll learn you're at the other world at which it has positive utility. Then gathering the evidence before choosing will lead you to choose whichever option is best at whatever world you're in. And that's better, in expectation, than picking whichever of the two options maximizes expected utility from the point of view of your

prior.

Example 5 (Anti-factivity)**.**

- $W = \{w_1, w_2, w_3, w_4\}$;
- $\mathcal{E}_{w_1} = \{w_2\}$ $\mathcal{E}_{w_2} = \{w_1\}$ $\mathcal{E}_{w_3} = \{w_4\}$ $\mathcal{E}_{w_4} = \{w_3\}$.

$C(-)$	w_1 1/4	w_2 1/4	w_3 1/4	w_4 1/4
o_1	1	2	-1	-1
o_2	0	0	0	0

Then:

- *C will pick o_1;*
- *If you gather the evidence, you'll pick o_1 in worlds w_1 and w_2, and o_2 in worlds w_3 and w_4, and doing that is exactly as good at worlds w_1 and w_2 as picking o_1, which C will do there, and it is better than picking o_1 in w_3 and w_4, which is again what C will do there, and so C prefers to gather the evidence.*

We can represent gathering evidence, updating on it, and picking in line with the updated credences as a new option on the list:

$C(-)$	w_1 1/4	w_2 1/4	w_3 1/4	w_4 1/4
o_1	1	2	-1	-1
o_2	0	0	0	0
GATHER	1	2	0	0

And, since GATHER *weakly dominates o_1 and o_2, C prefers it.*

Reflecting on this example gives us an interesting way to understand why learning evidence can be better, in expectation, than not learning it. Essentially, the availability of evidence makes available a new option in the decision problem that isn't there if you don't gather the evidence. In the example just given, the available options were zero-utility-for-sure or positive-utility-at-two-worlds-negative-utility-at-two-worlds. But, the evidence described there made available a different option: positive-utility-at-two-worlds-zero-utility-at-two-worlds. It made it available because by choosing to gather the evidence and then decide, and knowing how you'll update and then choose, you are essentially choosing the option whose utility at a world is the utility of whatever option you'll choose if you first update on the evidence at that world and then choose using those credences. And in the case just described, the option is at

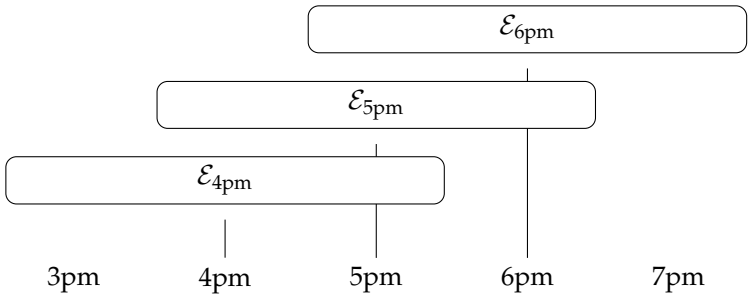


Figure 2.3: Partial illustration of the evidence function in the unmarked clock case (Example C).

least as good as each of the original options at every world and better than each at some. So it is better in expectation.

And, reflecting on this fact, we see that gathering evidence will have no value if the range of available options is rich enough. Suppose a coin has been tossed twice, giving four possible outcomes, HH, HT, TH, TT. And suppose you have the opportunity to learn whether the first coin landed heads—i.e. the disjunction HH or HT—or whether the first coin landed tails—i.e. the disjunction TH or TT. And suppose there are four possible options, which I'll write as the quadruple of their payouts at the four different worlds: (u_1, u_2, u_3, u_4), (u_1, u_2, v_3, v_4), (v_1, v_2, u_3, u_4), and (v_1, v_2, v_3, v_4). Then whatever my credences and whatever the values of these utilities, I will not pay to gain the evidence. For suppose my prior prefers (u_1, u_2, u_3, u_4) to the rest, and my posterior upon learning the first toss landed heads also prefers (u_1, u_2, u_3, u_4), while my posterior upon learning the first toss landed tails prefers (v_1, v_2, v_3, v_4). Then, if I were to prefer learning the evidence to avoiding it, I must prefer (u_1, u_2, v_3, v_4) to (u_1, u_2, u_3, u_4), since the former gives the pragmatic utilities of learning the evidence. But I don't have that preference, and I know I don't because that first option (u_1, u_2, v_3, v_4) is available to me, and I don't prefer it. And similarly for the other possibilities.

C. Williamson's unmarked clock. Externalists often contend that our evidence is not luminous to us; that is, we can have evidence that does not rule out our evidence being different from how it actually is. A neat illustration is Tim Williamson's (2014) example of the unmarked clock. You want to know the time. You can walk through to the next room and look at a clock. But alas, it is a fashionable clock of the minimalist sort favoured at the moment, and it has no numbers marked on it. It has only a sweeping single hand. Before you enter the room, you know the hand is pointing at 1 or 2 or 3 or ... or 10 or 11 or 12, but you know don't know which. If it's 1, your evidence will be that it's 12, 1, or 2; if it's 2, your evidence will be that it's 1, 2, or 3; and so on. That's how well your eyesight can discern the differences. Figure 2.3 illustrates part of the evidence function in this case.

Should you look at the clock? As in the example of corporate fraud above (Example B), it very much depends on your priors and the decision you'll

face. Suppose you currently assign equal credence to each possible position of the sweeping hand. The first option available pays a million dollars if it's at 1 or 2 or ... or 6, and nothing otherwise; the second pays a million dollars if it's at 7 or 8 or ... or 12, and nothing otherwise. Then you should gather the evidence—you'll do better in expectation if you do. But, as Nilanjan Das (2023) notes, if the first option pays a million dollars if the hand points to an even number and the second pays a million dollars if it points to an odd number, then you shouldn't gather the evidence, since the evidence is misleading about whether the time is even or odd: if it's even, you'll become twice as confident it's odd as you are that it's even, and vice versa.

It is worth pausing here to consider a philosophical use to which Simon Huttegger (2013, 2014) wishes to put the value of information framework. Picking up a suggestion by Brian Skyrms (1990a), Huttegger wishes to suggest that the value of information property can distinguish genuine learning events from events that change our credences but not via learning. So, suppose you know that after some event your credences will be given by either this credence function or that credence function, but you know neither which it will be nor what the mechanism will be by which you'll come to have them—Skyrms and Huttegger describe the credal change event as a black-box. Then, they say, the event is a genuine learning experience just in case, for every decision problem you might face, you'll prefer to have the experience first and then choose using the credences you'll end up with rather than choose using your current credences. So the reason that the change in your credences that result from, say, taking a hallucinogenic drug does not count as a genuine learning experience is that, from the point of view of your current credences, you'd expected yourself to choose more poorly after the change than you expect yourself to choose now.

It seems to me that this gives an elegant criterion by which to distinguish certain sorts of credal change that we don't want to describe as genuine learning events from those we do—events like taking a hallucinogenic drug or getting hit on the head, for instance. But I worry that it sets too high a standard for a genuine learning experience. For instance, in the case of the unmarked clock, it seems that we do have a genuine learning experience: my evidence will be factive, and it will rule out quite a lot of options that I currently rule in, and so it's pretty informative. But nonetheless it fails Skyrms' and Huttegger's test.

#5: Assessing updating plans. Throughout, we have assumed that, whatever evidence we gather, we update on it in the Bayesian's standard way by conditioning on the proposition learned. But, as Peter M. Brown (1976) showed, we can use the Hosiasson-Blackwell-Good framework to argue for this Bayesian assumption, at least in those cases to which the Value of Information Theorem originally applied, namely, cases of factive and partitional evidence.

An *updating plan* is a function that takes a possible world and returns a credence function. You might think of the credence function as the one the plan endorses at that world. We can easily define the pragmatic utility, at

a world, of an updating plan relative to a decision problem and tie-breaker function, to be the pragmatic utility, at that world, of the credence function it endorses at that world. Of course, what we'd most love is to follow the plan that takes each world to its omniscient credence function, that is, the one that assigns maximal credence to all truths and minimal credence to all falsehoods. But following that plan isn't available to us. Rather, we must pick a plan that assigns to two worlds the same credence function whenever those two worlds give rise to the same evidence. We'll call these the *available updating plans* relative to an evidence-gathering episode. Now, given an evidence-gathering episode, we can then ask which of the available updating plans has the greater expected pragmatic utility from the point of view of a prior credence function. Brown shows that, if the episode is factive and partitional, then updating plans that require you to condition on whatever evidence you learn maximize expected pragmatic utility.

What about cases in which the evidence is not factive or not partitional? Then Miriam Schoenfield (2017) shows that the updating plans that maximize expected pragmatic utility are not those that require you to condition on your evidence, but those that require you to condition on the fact you received the evidence you did. That is, in the unmarked clock case described above, if the sweeping hand points to 2 and I receive the evidence that it points to 1 or 2 or 3, I should conditionalize not on this evidence, but on the fact I received it, which is true only at the world at which it points at 2.

There is an interesting ongoing debate whether such an updating plan is really available to me (Carr, 2021; Gallow, 2021; Isaacs & Russell, 2023; Schultheis, ta). You might think it is not, since I could only implement it by reflecting on the evidence I in fact have, and inferring the worlds at which I would receive that evidence. But of course, in the cases at hand, I'm supposed not to know what evidence I have, and so presumably I can't reflect on it. However, externalists do think I should update by conditionalizing on the evidence I have, and you might think that this equally requires me to know what evidence I have.[4]

I won't delve deeper into this debate here, but I will note briefly that Gallow (2021) and Salow (2018) both argue that it should not be possible to enter yourself into evidential situations in which updating in the rationally required way is guaranteed to increase your credence in a false proposition. And yet that is exactly what happens in the unmarked clock case with respect to the proposition that the clock is pointing at an even number. This leads Salow to say that such evidential situations can't arise, offering an internalist account of evidence instead, and it leads Gallow to offer an alternative updating rule. My own take on this situation is that the evidential situation can arise and that you should update by conditionalizing when you face it, but the possibility of self-delusion to which this gives rise is not problematic because putting yourself in such a self-deluding evidential situation is rationally impermissi-

[4]Though Gallow (2021) describes an alternative updating rule the externalist might endorse instead.

ble from the point of view of the value of information: you simply shouldn't view the clock if it is the accuracy of your credence in the odd/even proposition that you care about; and you shouldn't do it if the decision you'll face with your credences is a bet on that same proposition.

To wrap up this section, it is worth noting that, if Schoenfield's updating rule is genuinely available to us, then the Value of Information Theorem holds for *any* evidence function, whether factive, partition, only one, only the other, or neither. That is, if we assume that we'll respond to evidence not by conditioning on the evidence we learn but on the fact that we learn it, then gathering evidence is always at least as good in expectation as not gathering it, and it is strictly better in expectation if you think learning it might lead you to choose a different option when you face the decision problem.

Brown's and Schoenfield's pragmatic arguments for updating

Some preliminary definitions:

- Given an evidence function $\mathcal{E}$, an updating plan R is *available in* $\mathcal{E}$ if, whenever $\mathcal{E}_w = \mathcal{E}_{w'}$, $R_w = R_{w'}$.

  That is, your updating plan can't discern the world any more precisely than your evidence can; it cannot respond differently at worlds at which you receive the same evidence.

- Given an evidence function $\mathcal{E}$ and a world w, let $\overline{\mathcal{E}_w}$ be the proposition that is true at all worlds at which your evidence is the same as it is at world w.

- Given an evidence function $\mathcal{E}$ and a prior C, an updating plan R is a Schoenfield plan for C and $\mathcal{E}$ if $R_w(-) = C(- \mid \overline{\mathcal{E}_w})$, whenever $C(\overline{\mathcal{E}_w}) > 0$.

Theorem 9 ((Brown, 1976; Schoenfield, 2017))**.** *Suppose $\mathcal{E}$ is an evidence function, C is a prior credence function and R, R' are updating plans. Then*

(i) *If R is a Schoenfield plan for C and $\mathcal{E}$, and R' is an available plan in $\mathcal{E}$, then, for any decision problem $\mathcal{D}$ and tie-breaker function τ,*

$$\text{Exp}_C(\text{PU}^{\mathcal{D},\tau}(R)) \geq \text{Exp}_C(\text{PU}^{\mathcal{D},\tau}(R'))$$

(ii) *If R is a Schoenfield plan for C and $\mathcal{E}$, and R' is an available plan in $\mathcal{E}$ that is not a Schoenfield plan for C and $\mathcal{E}$, there is a decision problem $\mathcal{D}$ and tie-breaker function τ such that*

$$\text{Exp}_C(\text{PU}^{\mathcal{D},\tau}(R)) > \text{Exp}_C(\text{PU}^{\mathcal{D},\tau}(R')).$$

Notice that, if $\mathcal{E}$ is factive and partitional, and R is a Schoenfield plan for C and $\mathcal{E}$, then $R_w(-) = C(- \mid \mathcal{E}_w)$, since $\mathcal{E}_w = \overline{\mathcal{E}_w}$.

Also notice that since the trivial updating plan on C, which takes every world and returns C, is always available, and it corresponds to not gathering the evidence at all, this theorem shows that, regardless of how your evidence function is, if you will update on new evidence using a Schoenfield rule, then gathering evidence is always at least as good in expectation as not, and it is strictly better, in expectation, if you think learning the evidence might lead you to change your mind about how to choose.

Chapter 3

The epistemic value of information

3.1 The epistemic value of gathering evidence

So far, we have focused on the pragmatic version of the Value of Information theorem, which tells us something about when you have *practical* reason to engage in a certain sort of evidence-gathering. But, as Graham Oddie (1997) showed, and Wayne Myrvold (2012) generalized, there is also a version that tells us something about when you have *epistemic* reason to gather evidence. Alejandro Pérez Carballo (2018) has extended Oddie's and Myrvold's approach in various ways, and we will see that Dorst et al. (2021) provide insights analogous to those they provided in the pragmatic case.

Recall: Hosiasson's insight is that the pragmatic value of a credence function is the utility of the option it leads you to choose, and the pragmatic value of an episode of evidence-gathering is the pragmatic value of the credence function it will lead you to have after you update your prior on the evidence you learn. But credence functions don't just have pragmatic value. We don't use them only to guide our decisions. We also use them to represent the world, and their purely epistemic value derives from how well they do that, regardless of whether we need them to help us choose.

Many ways of measuring this purely epistemic value have been proposed, but by far the most popular characterizations of the legitimate epistemic utility functions say that they are all *strictly proper*, where this means that, if we measure epistemic utility in this way, any probabilistic credence function expects itself to have strictly greater epistemic utility than it expects any alternative credence function to have; that is, it thinks of itself as uniquely best from the epistemic point of view; that is, it is epistemically immodest. Jim Joyce (2009) defends something close to this view, and Robbie Williams and I (2023) have recently argued for it in a different way. What's more, it is widely assumed throughout accuracy-first epistemology, which seeks to understand and ground the epistemic normativity of credences by investigating the optimal ways in which to pursue epistemic value understood as credal accuracy.[1]

[1]For an overview of this approach to epistemic normativity, see (Pettigrew, 2023b).

Strictly proper epistemic utility functions

An *epistemic utility function* EU takes a credence function C and a possible world w and returns $\mathrm{EU}(C, w)$, a real number or ∞ or $-\infty$, which measures the epistemic value of C at w.

Definition 3 (Strict propriety)**.** EU *is strictly proper if, for any probabilistic credence function P and any alternative credence function $C \neq P$,*

$$\mathrm{Exp}_P(\mathrm{EU}(P)) = \sum_{w \in W} P(w)\mathrm{EU}(P, w) > \sum_{w \in W} P(w)\mathrm{EU}(C, w) = \mathrm{Exp}_P(\mathrm{EU}(C))$$

Perhaps the most well-known strictly proper epistemic utility function is the so-called *Brier score*. Given a proposition, we say that the omniscient credence in it is 1 if it's true and 0 if it's false. The Brier score of a credence function at a world is then obtained by taking each proposition to which it assigns a credence, taking the difference between the credence it assigns to that proposition and the omniscient credence in that proposition at that world, squaring that difference, taking the average of these squared differences, and then subtracting the result from 1.

In the Brier score, each proposition is given equal weight in the average, but we can also give greater weight to some propositions than others in order to record that we consider them more important. This gives a *weighted Brier score*. This is important in the current context, since it allows us to explain why it is better, epistemically speaking, to engage in some evidence-gathering episodes rather than others, even when the latter will improve certain credences more than the former will improve others. The explanation is that the credences the latter will improve are less important to us. So, for instance, one evidence-gathering episode might, in expectation, greatly improve the accuracy of my credences concerning how many blades of grass there are on my neighbour's lawn, while another might, in expectation, only slightly improve the accuracy of my credences about the fundamental nature of reality, and yet I might favour the latter because the propositions it concerns are more important to me.

Another strictly proper epistemic utility function, less well-known but interesting nonetheless, is the *enhanced log score*. If a proposition is true, we score a credence in it by subtracting that credence from its own logarithm; if a proposition is false, we score a credence in it by subtracting that credence from zero. The enhanced log score of a credence function is then the average of these scores across all credences it assigns, and a weighted enhanced log score is a weighted average of them.

The Brier score and the enhanced log score

Definition 4 (Brier score). *The Brier score* $\text{Brier}(C, w)$ *of a credence function* C *at* w *is*

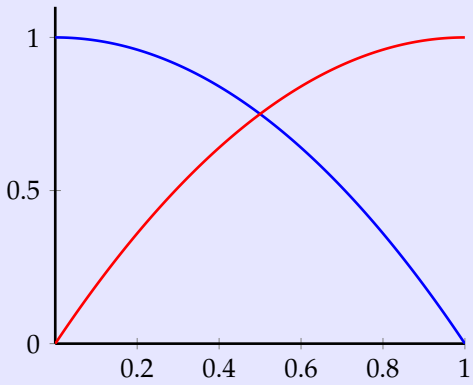
$$\text{Brier}(C, w) = 1 - \frac{1}{n} \sum_{X \in \mathcal{F}} |C(X) - V_w(X)|^2$$

where $V_w(X) = 1$ *if* X *is true at* w *and* $V_w(X) = 0$ *if* X *is false at* w*, and* n *is the number of propositions in* $\mathcal{F}$*.*

To give a weighted Brier score, we assign to each proposition X *in* $\mathcal{F}$ *a weight* $0 < \lambda_X < 1$*, where* $\sum_{X \in \mathcal{F}} \lambda_X = 1$*, and then define it as follows:*

$$\text{Brier}_\Lambda(C, w) = 1 - \sum_{X \in \mathcal{F}} \lambda_X |C(X) - V_w(X)|^2$$

The Brier score and any weighted Brier score are strictly proper.
In the diagram below, we plot the Brier score of a single credence in a true proposition in red (i.e. $y = 1 - (1 - x)^2$), and the Brier score of a single credence in a false proposition in blue (i.e. $y = 1 - x^2$).



Definition 5 (Enhanced log score). *The enhanced log score* $\text{Log}(C, w)$ *of a credence function* C *at* w *is*

$$\text{Log}(C, w) = \frac{1}{n} \sum_{X \in \mathcal{F}} V_w(X) \log(C(X)) - C(X)$$

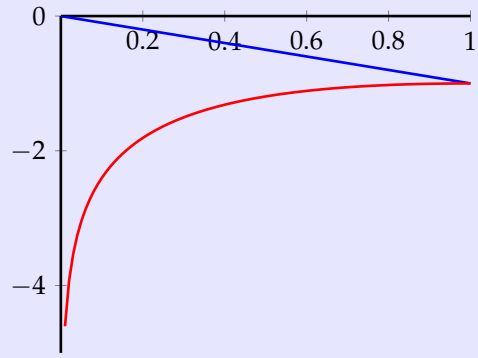
where again $V_w(X) = 1$ *if* X *is true at* w *and* $V_w(X) = 0$ *if* X *is false at* w*, and* n *is the number of propositions in* $\mathcal{F}$*.*

To give a weighted enhanced log score, we assign to each proposition X *in* $\mathcal{F}$ *a weight* $0 < \lambda_X < 1$*, where* $\sum_{X \in \mathcal{F}} \lambda_X = 1$*, and then define it as follows:*

$$\text{Log}_\Lambda(C, w) = \sum_{X \in \mathcal{F}} \lambda_X [V_w(X) \log(C(X)) - C(X)]$$

The enhanced log score and any weighted enhanced log score are strictly proper.

In the diagram below, we plot the enhanced log score of a single credence in a true proposition in red (i.e. $y = \log(x) - x$), and the enhanced log score of a single credence in a false proposition in blue (i.e. $y = -x$).



So now we have a way of assigning epistemic value to a credence function at a world. And so we can say that the epistemic value, at a world, of gathering evidence is the epistemic value of the credence function you'll end up with when you update on the evidence you'll get at that world—as before, we begin by assuming you update by conditioning on your evidence. And now we can state Oddie's epistemic version of the Value of Information Theorem: suppose you may gather evidence that will teach you which element of a particular partition is true, and suppose your epistemic utility function is strictly proper; then the expected epistemic value of gathering the evidence, from the point of view of your current credences, is always at least as great as the expected epistemic value of not gathering the evidence, from the same point of view; and, if you give some positive credence to a state of the world at which what you will learn will lead you to change your credences, then the expected epistemic value of gathering the evidence is strictly greater than the expected epistemic value of not doing so.

Oddie's Epistemic Value of Information Theorem

Theorem 10 ((Oddie, 1997))**.** *If* EU *is strictly proper and* $\mathcal{E}$ *is factive and*

partitional,

$$\mathrm{Exp}_C(\mathrm{EU}(\mathcal{E})) = \sum_{w \in W} C(w)\mathrm{EU}(C(- \mid \mathcal{E}_w), w) \geq \sum_{w \in W} C(w)\mathrm{EU}(C, w) = \mathrm{Exp}_C(\mathrm{EU}(C))$$

with strict inequality if there are w, w' such that $\mathcal{E}_w \neq \mathcal{E}_{w'}$ and $C(w), C(w') > 0$.

One thing that often surprises people about this result is that it seems to contradict the definition of strict propriety. According to strict propriety, every probabilistic credence function thinks it's best; but now we learn that it thinks that gathering evidence and updating on it to give different credence functions is even better. What's going on?

In fact, there is no contradiction: each probabilistic credence function thinks that it is better, in expectation, than any other specific credence function; but the updating plan isn't a specific credence function—it's different credence functions at different worlds. And strict propriety doesn't rule out a probabilistic credence function preferring a strategy that gives different credence functions at different worlds. Take, for example, the strategy, unavailable to all but God, of simply adopting, at a world, the omniscient credence function at that world, that is, the credence function that gives maximal credence (i.e. credence 1) to propositions that are true at that world and minimal credence (i.e. credence 0) to those that are false. Then this strategy gives the best credence function at each world. And so any credence function thinks of this strategy as better than itself, in expectation. But that doesn't contradict strict propriety.

As with the pragmatic version of the Value of Information Theorem, the reason Oddie's result holds is not that learning true evidence is guaranteed to improve your epistemic situation, and so certainly will improve it in expectation. As before, it's quite possible to acquire true evidence that is misleading. For instance, suppose my credence it's sunny is 10%, my credence it's windy is 40%, and my credence it's rainy is 50%. And suppose it's sunny. I then learn it's sunny or windy and my credence in sun becomes 20% and my credence in wind becomes 80%. Then, according to the Brier score, my epistemic utility dropped from 0.59333 to 0.57333. So my evidence was misleading and my epistemic situation deteriorated as a result of learning true evidence. But, as in the pragmatic case, Oddie's result holds because, in the particular conditions he places on the evidence-gathering episode, it will always be the case that any epistemic deterioration, once weighted by the prior's probability that it will happen, is outweighed by the epistemic improvements that are possible, once those are weighted by the prior's probability that they will happen instead.

3.2 Generalizing the epistemic Value of Information theorem

As with the Value of Information Theorem, we can generalize this result. As long as we set up an exchange rate between epistemic and pragmatic utility, we can factor in the cost of the evidence. That is, once we say how much pragmatic utility we're prepared to pay for a given amount of epistemic utility, we can say when gathering evidence is the right thing to do, rationally speaking. And, as before, we can use the expected epistemic utilities of different evidence-gathering episodes, with their costs factored in, to choose between them, and choose between them and doing something entirely different, which doesn't involve gathering evidence at all. And finally, we can generalize beyond factive and partitional evidence in a similar way.

Geanakoplos-style strengthening of Oddie's Theorem

Theorem 11 ((Dorst, 2020; Dorst et al., 2021; Levinstein, 2023))**.** *If $\mathcal{E}$ is factive, positively introspectible, and nested, then for any prior credence function C and any strictly proper epistemic utility function* EU*,*

$$\text{Exp}_C(\text{EU}(\mathcal{E})) \geq \text{Exp}_C(\text{PU}(C))$$

with strict inequality if there are w, w' such that $\mathcal{E}_w \neq \mathcal{E}_{w'}$ and $C(w), C(w') > 0$.

Simple Trust Principle and the epistemic value of information

Some preliminary definitions:

- An epistemic utility function EU is *additive and continuous* if there is a function $s : \{0,1\} \times [0,1] \to [-\infty, \infty]$ such that $s(1,x)$ and $s(0,x)$ are continuous functions of x, and

$$\text{EU}(C, w) = \sum_{X \in \mathcal{F}} s(V_w(X), C(X)).$$

- Given an updating plan R, a proposition X, and a real number $0 \leq t \leq 1$, let $\langle R(X) \geq t \rangle$ be the proposition that is true at all worlds w for which $R_w(X) \geq t$.

Simple Trust Principle Suppose C is your prior credence function and R is your updating plan R. Then, for any proposition X and any threshold t, the following should hold:

$$C(X \mid R(X) \geq t) \geq t.$$

Theorem 12 ((Levinstein, 2023))**.** *The following are equivalent:*

(i) *C, R satisfy the Simple Trust Principle*

(ii) *For any additive and continuous strictly proper epistemic utility function* EU*,*

$$\text{Exp}_C(\text{EU}(R)) \geq \text{Exp}_C(\text{EU}(C))$$

with strict inequality if there is w such that $R_w \neq C$ and $C(w) > 0$.

What about the cases we considered above?

A. Good and bad cases. In such a case, relative to any strictly proper scoring rule, this evidence will increase your epistemic utility in expectation. This is no surprise: in the bad world, your credence function will remain the same after learning the evidence; in the good world, it will become perfectly omniscient; and so in expectation, learning will be an improvement.

B. Misdirection vs complete information. In this case, relative to the Brier score, there are priors that will expect this information to increase epistemic utility and priors that will expect it to decrease. If your prior credence that the CEO did it is high enough, then there is no credences for CFO and COO that would making learning desirable; if your prior credence that the CEO did it is low (below 1/2, say), then learning can be desirable if your credences in CFO and COO are unequal enough.

Interestingly, relative to the enhanced log score, learning the evidence is never desirable. The reason is that, at the world at which it's the CEO, your credence function will assign credence zero to the true possibility, and this has epistemic utility $-\infty$ (since the logarithm of zero is negative infinity), and so the expected epistemic value of gathering the evidence is $-\infty$, whereas for any credence function its expected epistemic utility by its own lights is always greater than $-\infty$.

C. Williamson's unmarked clock. Again, relative to the Brier score, there are priors that will expect this information to increase epistemic utility and priors that will expect it to decrease. For instance, if you have equal credence in each of 1 or 2 or ... or 11 or 12, then you increase your Brier score in expectation by gathering the information, while if you lump nearly all of your credence onto one of the numbers, you decrease it. And this time, the same is true for the enhanced log score.

3.3 Assessing updating plans

As in the pragmatic case, we can appeal to measures of epistemic value to assess updating plans; and, when we do, we get the same results that we got in the pragmatic case. If your evidence function is factive and partitional, and your epistemic utility function is strictly proper, you maximize expected epistemic utility by choosing to update by conditionalization (Greaves & Wallace,

2006). And, more generally, regardless of what your evidence function is like, the available updating plan that maximizes expected utility relative to any strictly proper epistemic utility function is the one that tells you to condition not on your evidence but on the fact that you learned that evidence; that is, you maximize expected epistemic utility by choosing to update by Schoenfield conditionalization (Schoenfield, 2017).

Greaves & Wallace's and Schoenfield's epistemic arguments for updating

Theorem 13 ((Greaves & Wallace, 2006; Schoenfield, 2017))**.** *Suppose $\mathcal{E}$ is an evidence function, C is a prior credence function and R, R' are updating plans. Then:*

(i) *If R is a Schoenfield plan for C and $\mathcal{E}$, and R' is an available plan in $\mathcal{E}$, then, for any strictly proper epistemic utility function* EU*,*

$$\text{Exp}_C(\text{EU}(R)) \geq \text{Exp}_C(\text{EU}(R'))$$

(ii) *If R is a Schoenfield plan for C and $\mathcal{E}$, and R' is an available plan in $\mathcal{E}$ that is not a Schoenfield plan for C and $\mathcal{E}$, then, for any strictly proper epistemic utility function* EU*,*

$$\text{Exp}_C(\text{EU}(R)) > \text{Exp}_C(\text{EU}(R')).$$

Recall: If $\mathcal{E}$ is factive and partitional, then the Schoenfield plans are precisely the conditionalization plans.

3.4 Combining the pragmatic and epistemic values of information

I conclude this tour of the arguments and results concerning the value of information by noting that, once you set your exchange rate between the epistemic and the pragmatic, you can specify the all-things-considered value of an evidence-gathering episode. So, for instance, suppose your epistemic utility function is the Brier score; and suppose you fix a particular scale on which you'll measure your pragmatic utility (for recall that pragmatic utility is equally well measured on any of an infinite collection of scales, each obtained from any other by a positive linear transformation). Then, in order to specify the all-things-considered value of an evidence-gathered episode, you need to know how much pragmatic utility you consider equal to, say, 0.1 change in Brier score. If it's 0.1 units of pragmatic utility (measured on the scale we fixed), then your all-things-considered utility is just the sum of

your pragmatic utility (on the fixed scale) and the Brier score. If it's 0.2, then your all-things-considered utility is your pragmatic utility added to double the Brier score—0.1 change in Brier score is worth twice a 0.1 change in pragmatic utility. And so on.

This all-things-considered utility allows us to incorporate both the practical value of having credences as well as their representational value. Both are important to us, and the trade-off may well be important in certain cases. There will be cases in which learning the evidence increases our Brier score in expectation, but decreases the pragmatic utility of our credences because of the decision problem we will face with them. For instance, in the case of Williamson's unmarked clock, if we have equal prior credences in 1 or 2 or ...or 11 or 12, then we increase our Brier score by looking at the clock, but decrease the pragmatic utility of our credences for sure if the bet we're going to face is whether the hand points at an odd or even number. In that case, then, we must determine our exchange rate to discover whether looking at the clock is all-things-considered the right thing to do.

Part II

The epistemology of inquiry

Chapter 4

When should we inquire?

In the second part of this essay, I turn from the formal results and arguments that extend Hosiasson's and Good's Value of Information Theorem and focus on the literature on the epistemology of inquiry that has been developing recently in mainstream epistemology. My plan is to apply the insights from the first part to see what light they might shed on some of the central questions that have arisen in that literature.

The ambition of this part is imperialistic or, in Eric Schliesser's (ta) happier terminology, totalizing. In Schliesser's conception of synthetic philosophy, we take a totalizing approach when we take a formal framework from one area and apply it across a very wide range of cases in that area and in other areas: it might be the framework of natural selection, for instance, or the Bayesian treatment of uncertain belief. In my case, the framework is our standard theory of rational decision-making under uncertainty, developed in philosophy, economics, and statistics during the twentieth century.

Applied to the normativity of inquiry, this framework results in the value of information approach I described in the first part of this essay. Decisions whether or not to inquire are just a particular species of decision under uncertainty, and so they fall within the ambit of rational choice theory. To apply this theory to this sort of decision, we need an account of the value of a particular zetetic act, such as an evidence-gathering episode or a sequence of them. In the pragmatic case, this is given by Hosiasson's insight that the value of gathering evidence when you'll face a particular decision problem is just the value of the option you'll choose from those available in that decision problem with the credences you'll have after gathering the evidence. In the epistemic case, as Oddie teaches us, it's just the epistemic value of the credences you'll have after gathering the evidence. With these accounts in hand, we can apply rational choice theory to zetetic decisions just as we apply it to other decisions under uncertainty. And we can use it to say when it is pragmatically or epistemically or all-things-considered rational to start, continue, conclude, and reopen inquiry.

The totalizer claims, furthermore, that this approach says everything there

is to say about the rationality or correctness or appropriateness of doing these things. Rational choice theory is, after all, supposed to be a complete theory of the rationality of such choices.

To see the approach in action, let's ask four of the central questions in the epistemology of inquiry: when should we embark on a particular inquiry? when should we continue to pursue an inquiry on which we've already embarked? when should we conclude one? when should we reopen an inquiry we previously concluded?

4.1 What are the acts of inquiry?

When we ask these questions, we immediately face a further question that always arises when we use rational choice theory: how extended are the different actions that are available to the decision-maker, in this case, the potential inquirer? Take the example of a detective on a murder case trying to decide whether or not to inquire. Can they choose now to undertake the full inquiry, an extended action that involves first viewing the crime scene, then running the forensic tests, then taking the witness and suspect statements, then searching the suspects' homes? This is an action that might take a couple of weeks to complete: is it possible for them, at one point in time, to choose to perform that whole act? Or are they choosing to undertake only the first of these evidence-gathering episodes, namely, viewing the crime scene, but in the knowledge that taking that first step and learning its outcome will affect whether they continue with that inquiry or whether they move to another, or to something else entirely? Or are they choosing an even less extended action, such as walking towards the crime scene, or taking the first step in that walk?

Decision theorists have various ways of determining which actions are actually available to a decision-maker at the time of the decision. You might feel that the least extended action—the half-second action of taking the first step—is clearly available; perhaps the more extended action of viewing the crime scene—which we might say will take fifteen minutes—is less obviously available, but nonetheless is sufficiently available; and the fully extended, fortnight-long action of carrying out all of the inquiry is not sufficiently available to the potential inquirer at the very first point of the inquiry. One way to make this intuition clear is to say that the degree of availability of an action is the probability that the decision-maker would fully enact it were they to choose to do so. If the detective were to decide to walk towards the crime scene, it's very likely they would undertake and complete that full act; if they were to decide to view the crime scene, it's a bit less likely, but still very likely they would undertake and complete it; but if they were to decide to undertake the whole action, there are so many things that might knock them off course or force them to reconsider before the action is complete, that it is not sufficiently likely they would undertake and complete it.

For our purposes in what follows, we will choose the grain at which we describe the available actions to suit the examples we're considering. When

we're not interested in whether the individual might be knocked off course even in the execution of a single evidence-gathering episode, we'll simply take the whole episode to be an available action. But sometimes we are interested in this, as indeed we will be when we consider Julia Staffel's theory of transitional attitudes in Chapter 5; and in those cases we'll take less extended actions to be the ones available to us.

Whatever we choose, when we ask whether we should embark on an inquiry, we are asking whether we should undertake the first available action in that inquiry, which might be an evidence-gathering episode or it might be something shorter, such as an initial step in such an episode. And when we evaluate such an action, we must factor into the decision how likely it is that we'll proceed to the next step of the inquiry should we embark on the first, how much utility the outcome of that second step would have in conjunction with the outcome of the first, how likely it is we'll proceed to the next, as well as its utility, and indeed how likely it is that we'll complete the inquiry and what would be the utility of the outcome of doing that. So the options available to the potential inquirer are not *undertake-and-complete-the-inquiry* and *do-not-undertake-and-complete-the-inquiry*; they are usually *undertake-the-first-evidence-gathering-episode-of-the-inquiry* and *do-not-undertake-the-first-evidence-gathering-episode-of-the-inquiry*, where of course the second option might split further into many other possibilities, such as pursuing a different inquiry or doing something else entirely.

4.2 When should we begin, continue, or restart a particular inquiry?

So, how should you choose whether or not to begin an inquiry, or continue pursuing one you've already begun, or reopen one you previously closed? The totalizing answer is this: in the way you should choose everything else! From the pragmatic point of view, you should do so if doing so maximizes subjective expected pragmatic utility; from the epistemic point of view, you should do so if doing so maximizes subjective expected epistemic utility; and from the all-things-considered point of view, you should do so if doing so maximizes subjective expected all-things-considered utility.

An interesting consequence of this that we'll explore further below, again when we consider Julia Staffel's theory (Chapter 5): since an inquiry is a series of evidence-gathering episodes, it can be rational to embark on it even if not all of the episodes that make it up lead to improvements in the expected pragmatic or epistemic value of your credences, and even if some of the episodes lead to a decrease in those expected values, just as it can be rational to embark on a series of dental procedures even though you know that some of the individual procedures in the series will make things worse. Provided you're confident enough that you'll see the series through to the end, and provided the series in full leads to sufficiently great improvements in expectation, and provided your dental situation wouldn't be too much worse if the series got

interrupted in the middle, it is rational to embark on it (cf. Staffel's example of the detective and my example of the coloured necktie described below).

Sometimes, we embark again on an inquiry we have already completed: we double-check our results. On the face of it, this seems a puzzling practice. After all, we've undertaken the inquiry, and we've concluded it to our satisfaction. Why, then, are we undertaking it again? Woodard (2022) addresses this challenge in the knowledge framework, where the worry takes this form: if you have concluded an inquiry into a question, then you have come to know the answer, but you should not embark on an inquiry into a question unless you are ignorant of the answer, and so you should not inquire twice into the same question. She maintains in the face of this objection that it is permissible to double-check, and illustrates her point with a series of vignettes in which people appear to double-check rationally.

The Value of Information framework allows us to appreciate a number of circumstances in which it is rationally permissible, and perhaps even required to double-check:

In one sort of case, what we call double-checking is better described as acquiring a second sample of evidence from the same source. This is what we do if we check the door is locked when we leave our building, but then go back to double-check a minute later (cf. Woodard's example of Deming). We are not carrying out the same evidence-gathering episode twice; rather, we are carrying out two different but very closely related evidence-gathering episodes: Does the door open when I pull the handle this time? Does the door open when I pull the handle this different time? The first episode will already make me very confident the door is locked. What's more, the outcomes of the two episodes are very highly correlated with one another: that is, the probability of the door not opening when I pull it a second time, given it didn't open when I pulled it the first time, is extremely high and much higher that the unconditional probability of the door not opening on the second try. And so the second evidence-gathering episode can't hope to shift my credence that it's locked up upwards by more than a tiny amount. However, if I take the stakes to be very high and the cost very low, it might nonetheless be rational.

In another sort of case, we double-check something for which we had very good evidence at some point in the past, but not because we want to get even better evidence, but because we want to regain the confidence given us by the very good evidence we previously had. This is what we do, for instance, if we packed our favourite condiment in our luggage three days ago, but then today we open the luggage again to double-check it's there (cf. Woodard's example of Sam). In this case, we had very good evidence that the hot sauce is in our luggage three days ago when we put it there. Our immediate perceptual experience provided that evidence. Perhaps it warranted a credence of 99%. But gradually, as this experience has moved from being immediate to being a memory, it has come to support only a lower credence by today. Perhaps now the memory of that experience warrants only a credence of 90%. Perhaps I'm wrong about what I seem to perceive immediately before me only one time in a hundred, but I'm wrong about what I seem to remember perceiving imme-

diately before me one time in ten. And perhaps from the point of view of that lower credence, checking to see whether the hot sauce is there is something that maximizes expected utility, in which case double-checking is permitted, and maybe even required.

4.3 When should you cease inquiring further?

The totalizing view I am propounding here says that, just as you should embark on inquiry when doing so maximizes the subjective expectation of whatever variety of utility you're interested in, and similarly for continuing to pursue an inquiry and reopening a previously closed one, so you should cease inquiring further when continuing to inquire no longer maximizes that subjective expectation; that is, when there is an alternative available action that has greater subjective expected utility than continuing to inquire, whether that alternative available action is an evidence-gathering episode from a different inquiry, or some other action altogether, like making a sandwich.

Your reasons for gathering further evidence can just run out, and from that point of view it can be irrational to pursue your inquiry any further. For instance, this can happen if you care only about the pragmatic value of your credences as a guide to action in the face of a particular decision you know you'll face. At some point, you might come to know that all further evidence-gathering episodes that are actually available to you either won't change your mind about what to choose when faced with the decision, or that any that might change your mind are too costly. At this point, further inquiry is irrational from this myopic pragmatic point of view. While you might continue to improve your credences from an epistemic point of view, you achieve no further gains from a pragmatic point of view—at least none that issue from the decision you'll use those credences to face.

This can lead you to abandon before they're complete inquiries that it was nonetheless rational to embark on in the first place. For instance, this might happen because the costs of gathering further evidence in that inquiry has increased since you began the inquiry; or because the stakes of the decision you'll face using the credences you'll form have decreased; or because it becomes cheaper to inquire in a different way, a way that you thought would be too costly at the beginning of your inquiry, but which you have since learned is actually rather inexpensive—for instance, we can imagine a detective who is scouring through all CCTV footage over a 24 hour period because they thought that DNA testing would be expensive, but who has recently learned it's very cheap and so switches to that, abandoning their original inquiry.

From the epistemic point of view, things are a little different. Unless you somehow acquire certainty about the correct answer to the question at which your inquiry aims, there will always be some evidence-gathering episode that you'll expect to improve your credence function from a purely epistemic point of view, though of course that episode may not be available to you, or it might be too costly, and those are reasons to cease that inquiry.

Now, you will rarely acquire the sort of certainty that concludes inquiry regardless of what actions are available or what cost they have. After all, for most inquiries, the evidence-gathering episodes don't give definitive answers to the target question. They give definitive answers to related questions that bear on the target question, such as when I gather evidence about what the weather forecast says as part of my inquiry into whether or not it will rain tomorrow.

This vindicates a point raised by Avery Archer (13) and Christopher Willard-Kyle (forthcoming), who argue that, in inquiry, there will nearly always be room for improvement from an epistemic point of view. They are responding to those who say that knowledge is the aim of inquiry, and that an inquiry concludes once the inquirer knows the answer to the defining question. They argue this can't be right because, even after you've achieved this knowledge, it's always possible to improve your epistemic situation. After all, you might obtain *better* knowledge of the correct answer: you might obtain a safer belief, even though your current belief is sufficiently safe to count as knowledge; or you might obtain the belief you currently have, but using an even more reliable process, even though your current belief was formed by a sufficiently reliable process; and so on.

One interesting possibility that this throws up is that even those who think it is knowledge and not mere accuracy that we value will need to provide something like a numerical measure of the epistemic value of a doxastic state; to wit, an epistemic utility function. After all, one upshot of Willard-Kyle's point is that someone who knows the answer to the question at which their inquiry is aimed must decide whether or not to continue to pursue this inquiry. As he points out, by doing so, they can continue to improve their epistemic situation, but presumably there are diminishing marginal returns from such efforts, and so they must weigh those expected gains against the expected gains brought by some other pursuit. So, for instance, the detective who now knows that the suspect was at the scene of the crime can continue to inquire about that in order to improve the quality of her knowledge of it, or she can turn her attention to another question, such as whether they have a motive. To choose between these two courses of action, she must be able to weigh the improvements that each will bring in expectation. And that requires some way of measuring their epistemic value. I won't pursue this any further.

The main claim of the totalizing view, then, is that you should treat zetetic decisions in the same way you treat other decisions: you may do them when they are among the options that maximize subjective expected utility, you should do them when only they do this, and you should refrain from doing them when they don't. They are, after all, simply choices to do certain things; and their outcomes can be evaluated for their utility in exactly the same way the outcomes of other choices can; and our uncertainty about which outcome will eventuate can be treated as it can for other decisions. It is true that there is a purely epistemic perspective from which we sometimes wish to assess an evidence-gathering episode, but we saw in Chapter 3 that we can accommo-

date that as well; and we can combine it with the pragmatic perspective to give the all-things-considered perspective. With all of this in hand, we can now turn to some of the views that have been developed in the recent literature on the epistemology of inquiry.

Chapter 5

Do we have transitional attitudes?

Recently, Julia Staffel (2019, 2021a,b, ta) has drawn an interesting distinction between what she calls *transitional* and *terminal attitudes*. On her account, during the course of an inquiry, we form transitional versions of the attitudes we seek, whether these are outright beliefs or precise credences. Only when the inquiry is complete do we form terminal versions of those attitudes. So, for instance, a detective who is methodically working her way through the body of evidence her team has amassed for her forms transitional credences concerning the identity of the culprit, and only after she has surveyed all this evidence does she form terminal credences on that matter. Staffel says that what distinguishes these attitudes is at least partly what we're prepared to do with them:

(i) we are prepared to act on terminal attitudes but not on transitional ones;

(ii) we are prepared to make assertions based on terminal attitudes but not on transitional ones;

(iii) we are prepared to feed terminal attitudes into future deliberation and reasoning;

(iv) we are not prepared to do any of these things with transitional attitudes.

In the section, I want to explore how the value of information framework from the first part of this essay can shed light on this distinction.

5.1 Transitional attitudes in empirical inquiry

Let me begin in a surprising way by offering an argument that there can be no transitional attitudes that answer to Staffel's description. Let's suppose I face a decision in the midst of my inquiry that I expected to face at the end. Whereas I'd hoped to check four different weather apps before deciding what clothes to pack for my trip, I've only had the chance to check two before my travelling companion calls from the next room to hurry me to the train station.

It seems I have no choice but to choose using the credences I have at that point, which have been obtained from my credences at the beginning of the inquiry by updating on the evidence I've received during its course to date. After all, what else is available to me? Of course, there are my credences at the beginning of my inquiry. Should I use those instead?

The problem with that suggestion is two-fold. First, those credences themselves don't think I should use them, at least if the evidence-gathering episodes I've embarked on so far are ones for which embarking on them has greater pragmatic value in expectation than not embarking on them relative to my prior credences and relative to this decision problem. Sure, my prior credences would have liked it even more if I'd been able to complete my inquiry before facing this decision. But the world has prevented that and I must act now. And unless we have certainty about the relevant propositions, there will always be further evidence-gathering episodes we'd like to undertake, if they were cost-free and available to us, but often the world conspires to make them too costly or simply unavailable. So this predicament is not one in which we find ourselves only in the middle of an inquiry, but that doesn't prevent us from making decisions when we're less than certain.

The second problem with reverting to my prior credences to guide my decision is that, if I choose using those priors rather than the credences I have mid-inquiry, I will violate the Principle of Total Evidence. After all, while I might not have gathered all the evidence I wanted to gather before I had to face the decision, I did gather some, and choosing using my priors from the beginning of the inquiry is to ignore that evidence.

In the end, then, if I must act either on my priors or on the credences I hold mid-inquiry, I should act on the latter. But then they do not have the features that Staffel lists as distinctive of the transitional. What's more, the same goes for making assertions on the basis of those credences and feeding them into further deliberation and reasoning. Even if my inquiry is interrupted before it's complete, I should nonetheless choose whether and what to assert based on the credences I have reached by that point, and it is those credences I should feed into future deliberation and reasoning. To make an assertion is simply to undertake an action of a certain sort and its correctness can be evaluated in the same way as other actions by looking at its expected utility; so the analysis just given applies as much to the decision whether and what to assert as to other decisions. And surely when we feed attitudes into our further deliberation and reasoning processes, we want to satisfy the Principle of Total Evidence, and so it is again the credences we've reached mid-inquiry, before that inquiry was interrupted, that we should use.

In fact, this argument doesn't quite establish that there are *no* transitional attitudes of Staffel's sort. If an inquiry is made up of a series of evidence-gathering episodes, and each has positive expected pragmatic utility from the point of view of the credences you have just prior to undertaking it, then the argument suggests that no attitudes formed during that inquiry have the hallmark features of the transitional. But, as we saw above, not all inquiries are like that. Sometimes the whole sequence of evidence-gathering episodes is

such that we expect our credence function to be better after they're all completed, but there are points in the course of the investigation at which we expect that our credence function will be worse. This might happen, for instance, if we string together a bunch of Das' (2023) biased inquiries, where those in the first stretch are biased in one direction and those in the second are biased in the other, but taken together, they aren't biased in either direction.

For instance, take an example that Staffel considers in her recent book manuscript (ta). A detective asks her team to divide up the evidence they've gathered into that which suggests the first suspect is guilty and that which suggests the second suspect is guilty. She plans to begin by working through the first set, then moving on to the second set. In this case, while her prior credences expect the credences she'll have once she's worked through both sets to be better than they expect themselves to be, they also expect the credences she'll have once she's only worked through the first set to be worse that they expect themselves to be—after all, they think working through the first set is a biased inquiry, since it's guaranteed to raise her credence in the guilt of the first suspect. And so, if she's interrupted just as she completes the first set and suddenly has to make a decision she was hoping to make only at the end, she might well decide not to use her current credences. And in that sense they seem to be transitional.

In fact, I think this particular example doesn't quite work. The problem is that, at all points in the detective's inquiry, her evidence is luminous to her; that is, it is introspectively accessible to her. As she leafs through the evidence against the first suspect, she learns certain facts—their DNA was found at the scene, their alibi is shaky, and so on—but she also learns that she learns those facts. And so what she in fact learns is a proposition of the form *E and the strongest proposition I learned is E*. And so her evidence-gathering episodes are in fact factive and partitional: what she learns will be true, and it will be true that she's learned that thing, so the proposition *E and the strongest proposition I learned is E* will be true; what's more, for any two propositions E and E' that she might learn, the propositions *E and the strongest proposition I learned is E* and E' *and the strongest proposition I learned is* E' cannot both be true at the same world. And of course she knows in advance that the evidence will be this way. And so, by the original version of the Value of Information Theorem, gathering the evidence is always better in expectation than not gathering it. While it looks like the detective carries out a series of biased evidence-gathering episodes, she does not.

This is a lesson of Miriam Schoenfield's (2017) treatment of these cases. Sometimes, our evidence can look biased, or at least look as if it doesn't satisfy Geanokoplos' three conditions: factivity, positive introspectibility, and nestedness (recall from Section 2.6). For instance, if I am about to ask you what's in the fridge, and I know you'll say either 'Lettuce or cabbage' or 'Cabbage or radishes', then it appears that the propositions I might obtain as evidence are overlapping but not nested, and so there will be decision problems such that I shouldn't choose to obtain this evidence if I will face them afterwards—I will pay more for bets on whether there's a cabbage in the fridge after learning this

evidence than I would pay before, because my credence that there's a cabbage is guaranteed to rise whatever you tell me. But of course what I really learn if you say 'Lettuce and Cabbage' is not only that there is either lettuce or cabbage in the fridge, but also *that you said that*. And if you now consider the propositions *There is lettuce or cabbage in the fridge and you told me so* and *There is cabbage or radishes in the fridge and you told me so*, we see now that there is no overlap in the possible evidence I might learn: it is factive and partitional, and so, whatever decision I'll face, I should take that evidence if it doesn't cost too much.

So, in order to find a variety of inquiry in which there might be transitional credences, we must look to those cases in which the evidence we gather is not luminous to us, and so not necessarily partitional. Here is such a case:

D. Coloured headscarf. You have a headscarf in your pocket that is one of four colours: red, rose, peach, or orange. These colours are not very easily distinguishable. I have the opportunity to observe the necktie under two different lighting conditions. Under the first, if it's red or rose, I'll learn it's red or rose or peach; and if it's peach or orange, I'll learn it's rose or peach or orange. Under the second, if it's red or rose, I'll learn it's red or rose or orange; if it's peach or orange, I'll learn it's red or peach or orange. But, importantly, in both cases, I won't also learn that I learn these propositions; my evidence will not be luminous to me. The situation is illustrated in Figure 5.1. Then whichever prior credence function I have, the following are true:

(i) There is a decision problem such that viewing the necktie under the first lighting condition is pragmatically worse in expectation than not doing that if I will face that decision problem.

After all, that viewing will raise your probability in Rose or Peach for sure, and so the inquiry is biased and you violate the Weak Reflection Principle.

(ii) There is a decision problem such that viewing the necktie under the second lighting condition is pragmatically worse in expectation than not doing that if I will face that decision problem.

After all, that viewing will raise your probability in Red or Orange for sure, so again your inquiry is biased and you'll violate the Weak Reflection Principle.

(iii) For any decision problem, it is pragmatically better in expectation to view the necktie under both conditions than under neither if I will face that decision problem.

The combination gives an evidence function that satisfies Geanokoplos' conditions from Theorem 5—that is, it is factive, positively introspectible, and nested—and so the Value of Information Theorem holds of it.

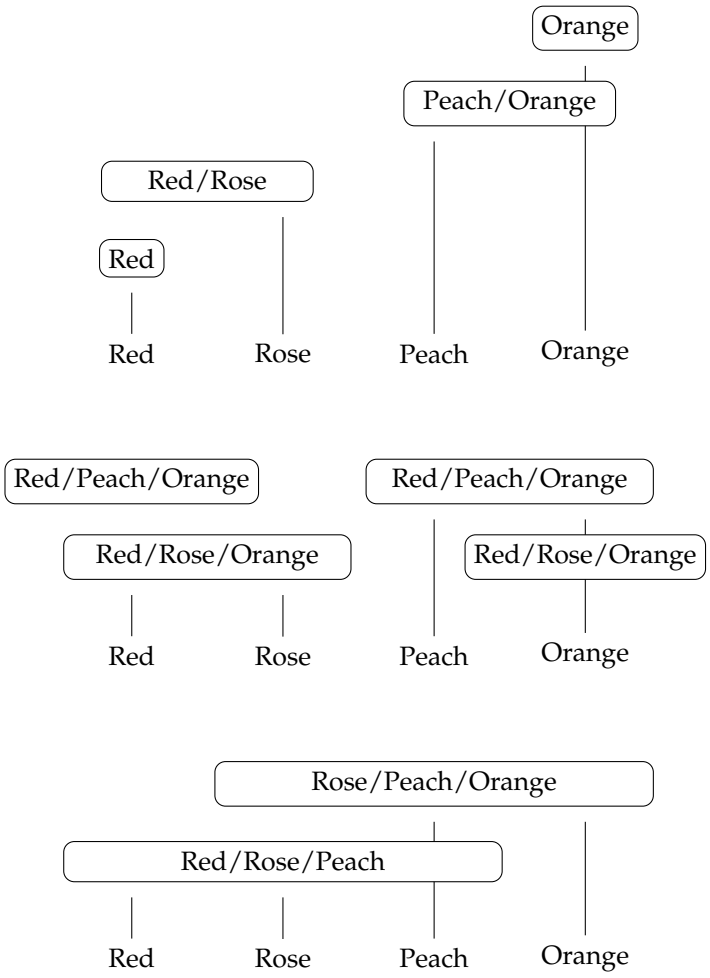


Figure 5.1: Illustration of the evidence functions in the red-orange necktie case. The evidence function for the first episode is at the bottom, then the evidence function for the second, and finally the evidence function for the combination is at the top.

So, do we here have a genuine case of transitional credences? Suppose I hope to view the necktie under both lighting conditions, and I'm almost certain I'll be able to before facing my decision. However, disaster strikes and I'm interrupted so that I must make the decision after only viewing it under the first lighting conditions. What's more, the decision I must face is one of the ones for which viewing under the first lighting condition is worse in expectation than not viewing. Let's suppose I face a bet on whether the necktie is rose or peach. Before viewing, I had credence 1/2 that it's rose or peach, because my prior was uniform across the four possibilities. After viewing, whether I learn it's red or rose or peach, on the one hand, or whether I learn it's rose or peach or orange, on the other, my credence that the necktie is rose or peach rises to 2/3. Now suppose I must choose whether to accept a bet that pays £10 if the necktie is rose or peach and £0 if it's red or orange, and which costs £6. So beforehand, I'd reject the bet as too expensive, but after viewing, whatever I learn I'll be confident enough that it's rose or peach to accept it.[1] So my priors would not have wished me to undertake only the first viewing before deciding; they'd have liked me to undertake both, and were confident enough I would that they went ahead with the inquiry. But now the first viewing has happened and I've been interrupted before the second viewing can take place. With which credences should I now approach the bet? My priors, which disapprove of my posteriors, because they'll take the bet? Or my posteriors, which disapprove of my priors, because they won't take the bet?

We might hope to appeal to the Principle of Total Evidence to adjudicate this case. It says that I should use my posteriors. But it was precisely the Principle of Total Evidence, or something very close to it, that both Hosiasson and Good were trying to justify using their Value of Information arguments. And of course those arguments don't apply to this case because the information gained does not have greater value in expectation from the point of view of the priors than the priors themselves have. So perhaps it's simply not true that, in the sort of situation we are considering, where the evidence is gathered as a result of a biased inquiry, we are rationally required to use the credences based on our total evidence to make decisions. Perhaps in such situations we are rationally permitted—perhaps even rationally required—to look to the credences we had before we acquired this evidence.

In this case, I think we should say that using my prior credences is rationally permissible; but using my posteriors is also rationally permissible. Why? Well, the fact that my evidence and credences are not luminous to me at the later time creates a strange situation in which there are really three vantage points from which I might assess the decision. First, there are my prior credences. They wouldn't take the bet at that price, and they'd prefer to use themselves to make the decision rather than to gain more evidence before deciding. Second, there are my posterior credences. I don't know what they are, but I know enough about them to be sure that they'd accept the bet at the

[1] Assuming here that utility is close enough to linear in money for these amounts.

stated price, and they'd prefer to use themselves to make the decision rather than use my priors. But third, there is the vantage point from which I reason consciously after acquiring the evidence. It's not very clear how to represent this vantage point, but it seems sufficiently uncommitted on the first-order question of the colour of the necktie to be able to appreciate the claims of both the prior and posterior standpoints. And from this more neutral standpoint, which is in any case the standpoint from which we must actually decide which credences to use, we can see the merits of both cases. So here, again, it is at least permissible to use as a guide to action and assertion and as an input into future reasoning the posterior credences that one has acquired in the middle of an inquiry that was interrupted before it was completed. So again we've failed to find transitional attitudes.

However, there are cases that seem to answer to Staffel's description. For recall Williamson's case of the unmarked clock. In that, the sweeping hand of an unmarked clock points at either 1 or 2 or ... or 11 or 12. If you view it and it points at 2, you'll learn it points at 1, 2, or 3; if you view it and it points at 3, you'll learn it points at 2, 3, or 4; and so on. Now suppose you will face a choice between Option A, which gives £10 if the hand points at an even number, and Option B, which gives £10 if the hand points at an odd number. And, as usual, you can decide whether or not to view the clock before making that choice. Then, if you prior is uniform across the twelve possibilities, it will prefer not to view the clock. It will expect itself to do better than the posteriors you'll obtain from such a viewing. But of course your posterior will prefer to use itself, rather than the prior.

So far, so similar to the previous example of viewing the necktie under different lighting conditions. But in the case of Williamson's clock, unlike in the case of the necktie, there is an external perspective from which to adjudicate the two claims. For, while your evidence and your resulting posterior credences are not luminous to you in this case—if they were, you could update on them and thereby learn exactly where the hand points—you do know the evidential situation that formed them. And one consequence of the evidential situation is this: (i) if the hand in fact points towards an even number, you'll become more confident that it points towards an odd number; (ii) if the hand in fact points towards an odd number, you'll become more confident it points towards an even number. And so you have evidence that, whatever you've learned, it will lead you to pick the wrong option out of Options A and B: if the hand points at an even number, it will lead you to choose Option B, which will then gain you nothing; if the hand points at an odd number, it will lead you to choose Option A, which will then gain you nothing. It is an interesting case in which, while your evidence is factive, it's anti-reliable about this particular question: it is guaranteed to make you strictly less accurate in your credences in the propositions concerning whether the hand points to an odd or even number. Indeed, whichever strictly proper epistemic utility function to measure the accuracy of our credences just in those two propositions—*The clock's hand points to an even number* and *The clock's hand points to an odd number*—learning this evidence is guaranteed to reduce it.

In this case, then, we have a genuinely transitional credence. It's quite possible that we could rationally choose to view the clock as the first step in an inquiry, even if the choice we'd ultimately face is between Options A and B. As long as the later steps of the inquiry rectify the problem posed by this earlier stage, it would be rational to embark on the inquiry, providing we are confident enough we'll be able to see it through to the end. However, if we are interrupted after the first stage, we should not be prepared to use the credences we then have to choose between Options A and B. Those credences are, then, genuinely transitional.

So there are credences that count as transitional on Staffel's definition. But it is notable how rare they are. Most inquiries don't include steps like the unmarked clock case. It is also notable that the credences' transitional nature is very sensitive to the decisions to be faced with them or the propositions over which they are defined. If instead we will face a bet on whether the hand points at a number between 1 and 6 or at a number between 7 and 12, then both prior and posterior credences will prefer to use the posteriors. And if we use the Brier score to measure the epistemic utility of the credences in all propositions of the form The clock's hand point to n, then again the posterior will be preferable, even from the point of view of the prior.

5.2 Transitional attitudes in logical reasoning

Staffel wishes to apply her account of transitional and terminal attitudes not only to cases of empirical inquiry, like the detective investigating a murder, but also to cases of logical reasoning and other *a priori* inquiry, like the logic student who uses truth tables to establish whether $(p \rightarrow (p \rightarrow p))$ is a tautology or not, the diner using mental arithmetic to calculate a restaurant bill, or the person reflecting on the evidence they've acquired to draw out its consequences. Following Ian Hacking (1967), as well as recent developments of his view by Robbie Williams (2018) and me (2020), we might appeal to the value of information approach to inquiry to model logical and *a priori* reasoning as well.

Hacking's idea is that, just as empirical evidence serves to rule out certain possible worlds as not actual, so the fruits of logical inquiry and other forms of *a priori* reasoning also serve to rule out worlds as not actual. Now, the worlds that they rule out are not possible worlds in the standard sense. When the logic student completes the first row of the truth table for $(p \rightarrow (p \rightarrow p))$ and thereby discovers that this formula is true when p is true, this rules out the world in which that formula is false when p is true. But this isn't a genuine possible world, because it isn't logically possible. Rather, Hacking suggests, it is a *personally possible world*: that is, it is possible from the point of view of the student's original epistemic position. According to Hacking, we might represent the student as distributing credences over these personally possible worlds in such a way that they sum to 1 and her credence in any proposition is the sum of her credences in the personally possible worlds at

which that proposition is true. In this way, we can give a sort of Bayesian representation of a reasoner who is not logically omniscient. And then we can understand logical learning in the same way the Bayesian understands empirical learning: both rule out worlds; and, having ruled out worlds, we assign zero credence to them and then scale up our credences in the remaining worlds so that they again sum to 1. On this picture, logical learning is formally represented exactly as empirical learning is, and so the Value of Information Theorem and Geanakoplos' and Dorst's generalizations, as well as Oddie's epistemic version of the theorem and its generalizations, all hold of logical learning just as they do of empirical learning.

For instance, take our logic student, who is considering the formula $(p \rightarrow (p \rightarrow p))$. Perhaps, prior to her inquiry, four worlds are epistemically possible for her: in the first, the formula is true when p is true and true when p is false; in the second, it's true when p is true but false when p is false; and so on. Perhaps she distributes her prior credences equally over these four personally possible worlds, giving $1/4$ to each. And so she has credence $1/4$ that the formula is a tautology, since that is true only at the world in which the formula is true when p is true and true when p is false. Then, when she constructs the first line of the truth table, she learns the formula is true when p is true, and rules out worlds at which it is false when p is true. She updates her credences, thereby coming to have credence $1/2$ that the formula is a tautology. And finally, when she learns it's true when p is false, she updates her credences again, becoming certain it's a tautology.

One appealing feature of this approach to logical and *a priori* reasoning, which Hacking already identified and I developed a little further, is that it furnishes us with an account of the rationality of credences in logical propositions on which it is rationally prohibited to be uncertain of simple logical truths, but rationally permitted to be uncertain of more complex ones. And indeed this account is built on the Value of Information Theorem. The point is that a simple logical truth is easily established, and so the cost of undertaking the logical inquiry needed to learn it is low, and yet it might well have a moderate pragmatic utility. And so if you don't incur the cost and learn it, you have irrationally failed to maximize your pragmatic utility in expectation. Complex logical truths, in contrast, are not easily establish; the cost of inquiring into them is high; and it is often not worth paying that cost, given the expected pragmatic pay off.

If this is the right way to represent logical ignorance and logical learning, then the same points apply to transitional attitudes within logical or other *a priori* reasoning that I made about such attitudes in the midst of empirical inquiry above. And indeed it might turn out that logical inquiries that give rise to genuinely transitional credences are even rarer than empirical inquiries that do. After all, many of the cases in which our evidence and credences are not luminous to us are ones in which that evidence is perceptual: viewing a necktie under certain lighting conditions; viewing an unmarked clock. And it isn't clear whether there are similar ways in which we might acquire logical evidence that would lead to the same failure of luminosity.

Chapter 6

When and how should we pay attention?

Let's turn now to an example that motivates some of Jane Friedman's writings about inquiry, which in turn launched the recent literature about inquiry in mainstream epistemology. I'll quote at length:

> I want to know how many windows the Chrysler Building in Manhattan has (say I'm in the window business). I decide that the best way to figure this out is to head down there myself and do a count. [...] Say it takes me an hour of focused work to get the count done and figure out how many windows that building has. [...] Now think about the hour during which I'm doing my counting. During that hour there are many other ways I could make epistemic gains. [...] First, I'm a typical epistemic subject and so I arrive at Grand Central with an extensive store of evidence: the body of total evidence, relevant to all sorts of topics and subject matters, that I've acquired over my lifetime. Second, I'm standing outside Grand Central Station for that hour and so the amount of perceptual information available to me is absolutely vast. [...] However, during my hour examining the Chrysler Building I barely do any of that. I need to get my count right, and to do that I really have to stay focused on the task. Given this, during that hour I don't extend my current stores of knowledge by drawing inferences that aren't relevant to my counting task, and I do my best to ignore everything else going on around me. And this seems to be exactly what I should be doing during that hour if I want to actually succeed in the inquiry I'm engaged in. [...] There is an important sense in which I succeed in inquiry by failing to respect my evidence for some stretch of time. It's not that my success in this case comes by believing things my evidence doesn't support, but it does come by ignoring a lot of my evidence and failing to come to know a great deal of what I'm in a position to know. (Friedman, 2020)

Friedman concludes that, in cases like this, there is a tension between epistemic norms and zetetic norms. Here are two of the epistemic norms Friedman mentions:

> **EP**$_a$ If one has excellent evidence for p at t, then one is permitted to judge p at t.
>
> **EP**$_o$ If one has excellent evidence for p at t, then one ought to judge p at t.

And here is the zetetic norm:

> **ZIP** If one wants to figure out [the answer to a particular question], then one ought to take the necessary means to figuring out [that answer].

Does the Value of Information approach allow us to say anything illuminating about this apparent tension? I think the natural thing to say here is that, as Friedman faces the Chrysler Building, she faces a choice between a number of different evidence-gathering episodes she might undertake. Some of them are the ones that form the inquiry she is there to undertake, namely, determining the number of windows in the building; some involve attending to sensory information and perhaps testimony that is available at the spot where she's ended up, but which is irrelevant to her inquiry; and some involve drawing inferences from the store of memories and other evidence she's previous collected, whether by deduction, induction, or abduction, which is again irrelevant to her inquiry.

Of course, it's rather unusual to think of these last episodes as involving evidence-gathering. After all, you already have the evidence, and you're simply drawing conclusions from it that you haven't drawn before. But, as I described in the previous section, I think it's reasonable to view logical and *a priori* reasoning as doing something similar to what gathering empirical evidence does. In both cases, they are ruling out states of the world that are in some sense possible.

So, having seen this, we can understand the logical reasoning that Friedman doesn't do when she's in front of the Chrysler Building as just another sort of evidence she doesn't gather, just as she doesn't gather the evidence she might gather if she were to attend to the conversation between the two commuters standing to her left, say. And once we do that, we can say that Friedman does the right thing by continuing with her window-counting inquiry so long as, at each stage, the evidence-gathering episode that comes next in that inquiry is the one that maximizes expected pragmatic or epistemic value among those episodes that are available to her. And if we see things in this way, there is no clash between an epistemic norm and a zetetic one once those norms have been correctly stated. There is just one norm: gather evidence in the way that maximizes expected utility, and respond to any evidence you gather by conditionalizing on it. And that norm governs what Friedman should do in front of the Chrysler Building. Reasoning that

attending to the conversation between the commuters is very unlikely either to increase the epistemic or pragmatic utility of her credences by much, and noting that continuing to count the Chrysler windows will increase the pragmatic utility of her credences for sure, and realising that she cannot do both, Friedman should continue to count.

It might seem that this all rather misses the point. In Friedman's example, you might think, the question is not whether to gather certain evidence that surrounds you as you stand outside Grand Central Station. The point is that you already have that evidence, whether you want it or not, simply by being there. You cannot help but have it. And so the real question is whether to incorporate it or not when doing so might take up resources that can be used to gathering the evidence about the number of windows. And that's what creates the clash between epistemic and zetetic norms, since Friedman thinks you should not incorporate the evidence because of how doing so interferes with your inquiry, but presumably standard epistemic norms—and, in particular, $\mathbf{EP}_o$—say you should incorporate that evidence, since you have it.

In the end, this comes down to when you want to say that someone *has* certain evidence, and what exactly that means. The fact that all this evidence is to hand around me outside Grand Central Station does not necessarily mean that I have it. It might just mean that it's easily accessible to me, should I wish to gather it. But even if we do wish to say that we have some of it because, perhaps, it's impossible to ignore certain things like the content of a very loud conversation or the presence of a very brightly coloured car nearby, we can model this in a way that allows us still to use the Value of Information framework. We can say that the evidence is now stored in us somewhere and somehow, but we haven't yet brought it to our attention; we haven't yet passed it to whatever part of our mind takes in evidence and alters credences in the light of it. Rather, it's sitting in storage waiting for us to decide whether or not to attend to it. But then the decision whether or not to attend to it is just like the decision whether or not to gather and update on some evidence, and the Value of Information framework applies. So again, I can know that there is the content of some loud conversation sitting in storage, and I can choose to bring it out of storage and attend to it or continue with my enumeration of the windows in the Chrysler building. And the totalizing Value of Information framework I've been proposing in this second half will tell me what to do.

It's natural to describe Friedman in front of the Chrysler building as choosing how to direct her attention: to the windows of the building or to the evidence from her surroundings. And the totalizing version of the Value of Information framework takes itself to have much to say about how we should direct attention. Like gathering evidence, directing attention is simply an action we perform—we might call it an epistemic action, since it is undertaken at least partly in order to alter our epistemic state. Indeed, gathering evidence is perhaps just one species of directing attention; another is extracting information from evidence already gathered by inferential processes, whether deductive, inductive, or abductive. And, as we saw in the discussion of logical and *a priori* inquiry above, the other species of directing attention are also epistemic

actions that we can evaluate using the Value of Information framework.

You might worry that we don't always choose how to direct our attention, and so it is not appropriate to represent us as making a decision between directing it in one way rather than another. But while its name suggests that decision theory is only about decisions we voluntarily and consciously make, and rational choice theory only about choices we voluntarily and consciously choose, there is no need to treat them like that. We can just as well use their tools to assess choices that are unconscious, involuntary, or both. We can say of these choices too that they maximize or fail to maximize expected utility. And, while we might hesitate to praise someone for doing something unconsciously and involuntarily that does maximize expected utility or blame someone for doing something unconsciously and involuntarily that fails to so maximize, nonetheless, we can say that it's better to do the former and worse to do the latter, and that whatever unconscious process is driving the involuntary action is serving the person who has it well in the first case and poorly in the second. That is to say: decision theory is a way of evaluating, from a particular point of view that includes credences and utilities, the selection of an option from a menu. Nothing in the theory requires that this selection is performed by a person; it might be performed by a non-human animal, an organism, a computer, a state, or an algorithm. And nothing in it requires that it be done intentionally or consciously.

This point also helps the totalizing view to accommodate Georgi Gardiner's (2022) observation that often we assess an individual's patterns of attention over a period, rather than individual acts of directing attention. We notice that people tend to pay attention to this rather than that. And we might notice that, while each individual instance within this pattern might be considered permissible when considered in isolation, the pattern itself is not. Gardiner argues that we should therefore take a more virtue-theoretic approach to our evaluation of a person's patterns of attention. But I think we should rather use the Value of Information approach I've been describing, but where the option to be assessed is not just an individual act of directing attention in one way rather than another, but instead a disposition to direct attention that is manifested over a period of time. This allows us assess Friedman's choice to continue directing her attention to the Chrysler windows, which is a single action; and it allows us to assess an individual's lifelong disposition to direct attention in particular ways, which is the sort of thing a virtue-theoretic approach will consider; but it also allows us to assess dispositions manifesting over shorter timescales than the virtue theorist will consider, such as when someone, during a week-long period of uncharacteristically heightened anxiety, focuses only on the possibilities that lead to mortal danger.

The Value of Information approach I've been proposing in this second part of the essay can also help us to understand the interesting phenomena of group attention that Gardiner discusses. She points out that groups can contain individuals who direct their attention in ways that are bad for them but end up helping the group, and there are groups in which all individuals direct their attention in good ways, but the group itself performs poorly. This

recalls what Mayo-Wilson et al. (2011) call *the independence thesis* in the philosophy of science. That thesis says that epistemically rational individuals might form epistemically irrational groups and epistemically irrational individuals might form epistemically rational groups.

Suppose, for instance, that each individual in a group begins with the same prior credences. And suppose that each directs her attention to just one feature of a situation, leading her to collect evidence in a very biased fashion that is, in expectation and from her individual point of view, worse than not collecting it. But suppose further that each individual directs attention to a *different* single feature of the situation. Then the group itself might end up with a very balanced and rich set of evidence, since there will be no overlap between the pieces of evidence gathered by the individuals. In this situation, each individual will end up with credences that have lower epistemic and pragmatic utility in expectation than they would have had if she had directed her attention in a less focussed way and gathered evidence more evenly. But, if we take the group's credences to be the shared prior credences updated on the group's total evidence—which is this balanced and rich set—then they will have higher epistemic and pragmatic utility in expectation than they would have had if each individual had gathered the evidence more evenly and ended up with a total body of evidence that was also very balanced, but this time much less rich, since there would be a lot of overlap.

In general, the teleological approach embodied in the Value of Information framework I've been describing is useful in non-ideal epistemology—its totalizing ambitions are not restricted to ideal epistemology. In that approach to epistemology, we look at different components of our epistemic practices—how we inquire, how we reason, how we direct attention, and so on. Each of these practices leads ultimately to the credences we have. So we can evaluate them by the expected epistemic or pragmatic utility of the credences they'll lead to. Non-ideal epistemology recognises that we are limited, and so the range of epistemic practices we can engage in is limited—Friedman's window counter can attend to the evidence about the windows or to the evidence about the couple's financial situation, but she can't do both; as I sit on a short train journey, I have the time to draw out some inferences from the evidence I received last night chatting to a friend and I have the time to read about my latest interest on Wikipedia, but I can't do both. And a teleological approach is well suited to that as well: when you assess a particular practice we engage in, you compare it only with others that are actually available to limited creatures like us. So you don't criticise the window counter for not attending to both sources of evidence, since this is simply beyond her capacities. While some of the main versions of this teleological approach to epistemology has often focused on general norms for ideal epistemic agents—betting arguments and accuracy arguments for Probabilism, Conditionalization, and so on—it is ideally suited to answering the sorts of questions that interest non-ideal epistemologists as well.

Chapter 7

Why should we not resist evidence?

In this final section, I want to ask what the Value of Information approach might say about what Mona Simion (2023) calls 'epistemic duties to believe', and particularly what it says about the sorts of violations of those norms that she gathers together under the heading of 'resistance to evidence'. The core of Simion's concern is that, in the past, epistemologists have based their assessment of an individual's doxastic state—the justification or rationality of a belief or credence—entirely on the evidence the individual in fact has, rather than basing it on both the evidence they have *and the evidence they should have had*. This means that the racist who simply resists any evidence that undermines their racist beliefs will count as rational and justified, as will the sexist who ignores evidence provided by a woman, or the climate denier who simply does not take on evidence contrary to their position. Simion seeks an epistemic duty that requires us not only to believe when we have collected evidence that supports a proposition, but also to believe when there was evidence at hand that supported that proposition, whether or not we in fact collected it. Here's the norm she gives:

> **DTB:** A subject S has an epistemic duty to form a belief that p if there is sufficient and undefeated evidence for S supporting p.

Of course, much is going to turn on what it means to say that *there is* sufficient and undefeated evidence, and Simion gives a detailed account of this. Using the Value of Information theorem, we might offer an alternative account: *there is* evidence available to an individual when the cost of gathering it would be very small, and certainly greatly outweighed by the expected utility of gathering it. (One hiccup here is that there might be very very many different pieces of evidence available where the cost of gathering each piece is very small, but we can't gather it all. But let's bracket those cases; the ones we consider here are not like that.)

So now let's consider the sort of case Simion has in mind and see what the Value of Information theorem tells us about them.

> **Case 1: Testimonial Injustice.** Anna is an extremely reliable testifier and an expert in the geography of Glasgow. She tells George that Glasgow Central is to the right. George believes women are not to be trusted, and therefore fails to form the corresponding belief. (Simion, 2023)

I think the subjective Bayesian's assessment of this case is a little different from Simion's, since they deal with the agent's subjective prior credences, while Simion works with a notion of evidential probability that many subjective Bayesians—and certainly this particular subjective Bayesian!—disavow. On perhaps the most natural subjective Bayesian reading, the case of George and Anna isn't a case of resistance to evidence, but rather a case of irrational priors. After all, let's take the evidence that George obtains in this situation to be that Anna says Glasgow Central is to the right. He might well incorporate that evidence exactly as the Bayesian says he should and yet retain a low or middling credence that Glasgow Central is to the right. For Simion says that George believes women aren't to be trusted, and so this is something that is encoded in the credence function he has when he meets Anna and hears her testimony. The Bayesian says he should conditionalize on his priors, but doing so will lead him to have something pretty close to his previous middling credence about the direction of Glasgow Central, since he'll think Anna's testimony is not much better than chance as an indicator of the truth. For the Bayesian, the situation is structurally akin to a case in which I irrationally believe that the thermometer on my wall is completely broken when in fact it's very accurate, and so when it tells me that the temperature is 20C, I update on that evidence exactly as the subjective Bayesian says I should, conditionalizing my priors on it, but it doesn't shift my credences, because my priors were irrationally inaccurate and treated the thermometer's reading are almost independent of the true temperature.

So, for the subjective Bayesian, George is certainly flawed, but it's not because he is resistant to the evidence Anna gives him; or, at least, it isn't because he's resistant to the evidence in the sense that he fails to incorporate it. It is rather because he has an irrational prior that leads him to have an irrational posterior after he does incorporate the evidence in the way his prior demands.

Of course, his irrational prior might be the result of having resisted evidence in the past. There are at least two ways George might have ended up with that prior:

On the first, his ur-prior, the credence function he has at the beginning of his epistemic life, might have assigned very low credence to the reliability of women's testimony, and that will be judged irrational since it's taking an extreme stand on a proposition about which George had no evidence at that time. What's more, if he assigns higher credence to the reliability of men's testimony, say, we will judge it further irrational because it differentiates between two cases when he has no evidence to justify such differential treatment.

On the second way he might have arrived at his irrational prior, his ur-prior might have assigned middling credence to the reliability of women's

testimony, just as it did to the reliability of everyone else's testimony, but then as he went through life he incorporated any evidence he received that told against women's reliability and failed to incorporate any evidence he received in its favour. This leaves him with the biased credence function he has when he meets Anna and hears her testimony. In this case, George exhibits genuine resistance to evidence he received, and Oddie's version of the Value of Information theorem tells us what went wrong with him: he failed to incorporate evidence when incorporating it would have improved his epistemic situation in expectation.

Let's turn now to a case raised by Simion in conversation:

> **Case 2: Climate change denier** Jon denies that there is an anthropogenic component to current dramatic changes in Earth's climate. Over the years, this has become such a large part of Jon's thinking that it constitutes part of his identity. A great deal of evidence to the contrary is available to him, but he resists it, perhaps unconsciously, because to face it and incorporate it properly would be to lose a belief that forms part of who he is; losing that belief would be very costly to Jon, resulting in anguish, disorientation, and alienating him from the epistemic bubble into which this belief has drawn him.

Surely, Simion contends, the Value of Information approach says that, in this case, Jon should not incorporate the evidence he has; and just as surely this is the wrong answer. Given the pain it will cause Jon to lose his belief that current climate change is entirely naturally caused, it almost certainly outweighs any expected pragmatic or epistemic utility he'll gain by gathering the evidence that will lead to this. So, even from his current point of view, where he assigns very high credence to the proposition that climate change is naturally caused, and therefore very low credence to the evidence he might gather changing his mind, the negative effects of changing his mind are so great that he still gives higher expected utility to not gathering the evidence. So, the Value of Information theorem says, he does nothing wrong by not gathering it. And that, Simion contends, is the wrong answer.

I think this is a case where it is helpful to distinguish two ways in which we might evaluate someone's actions: a purely subjective one, and a slightly more objective one. In the purely subjective sense, Jon indeed does nothing wrong: he best serves the ends and values that he actually has by avoiding the evidence that might change his mind about climate change; relative to his actual credences and actual utilities, this is what maximizes expected value. But in the more objective sense, we might criticize exactly those ends and values and so say that what he ought to do is not what maximizes in expectation the utilities that encode them. We might say that Jon ought not to have those values and the utilities that encode them; he ought not to have built his identity around an empirical belief in that way when doing so would make him so resistant to learning anything that would unseat it. If we do this, we might

instead evaluate him not from the point of view of his actual credences and actual utilities but from the point of view of his actual credences and the utilities *he should have instead*—utilities that do not place such negative value on coming to believe that climate change is anthropogenically caused. And once we do that, the Value of Information theorem delivers the result we want: Jon should gather the evidence.

Recognising these different ways in which we might evaluate someone's evidence-gathering behaviour reveals a flexibility in the Value of Information approach I haven't had cause to highlight so far. Formally, the approach needs a vantage point from which to assess an evidence-gathering episode, a representation of that evidence-gathering episode as an evidence function, and an account of how the individual would respond to the evidence should they receive it. In the pragmatic case, the vantage point consists in a probability function, a decision problem they will face, and a utility function: the probability function is typically taken to be the actual credences of the would-be evidence-gatherer at the point at which they have to decide whether or not to gather the evidence, and their utility function encodes their actual values at that point. And in the epistemic case, the vantage point consists in a probability function and an epistemic utility function. As the example of Jon shows, in the pragmatic case, sometimes we might want to use other utilities than the individual's own; perhaps their own are immoral or unreasonable or self-undermining or in some other way flawed. But equally we might want to use other probabilities than the individual's own credences. We might want to use evidential probabilities, if we think there are such things. Or we might want to use the credences the individual *ought to have had*, not the ones they actually have, where we take these to be the ones they'd have had had they acted correctly in the past. This is another interesting theme that emerges in Simion's work on resistance to evidence.

Think again of Jon, the climate science denialist. We said he decides not to gather evidence that might persuade him that climate chance is manmade because that risks something he values greatly, even though he shouldn't. But consider his friend, Jim, also a climate denialist, who faces the same decision whether to gather evidence that might overturn his scepticism. Unlike Jon, Jim's utilities are in order: he ascribes no value to being a denialist and cares only about discovering the truth. However, because of how he's directed his attention in the past—or had his attention directed by his social environment—he has only ever picked up on evidence against the anthropogenic origins of climate change and he's always ignored or missed evidence in its favour, even when it's been readily available at a low cost to him. Let us stipulate that, in these past cases, the Value of Information approach would judge that Jim acted irrationally in not gathering this latter body of evidence. His current credences, shaped by his past history of gathering and not gathering different pieces of evidence, assigns such a high credence to the natural origin of climate change, and so thinks it so unlikely that gathering further evidence will change his mind that it's rational *from that point of view* not to pay the costs of gathering it. Nonetheless, Simion thinks and I agree there is

a sense in which we want to say that Jim should gather it, just as Jon should. In this case, it isn't Jim's utilities we want to fix up, but his credences. It's not that we want to appeal to some objectively correct alternative credences, such as the evidential probabilities, but rather to the credences that Jim would have had had he gathered evidence in the past in the ways that were rationally required by his own credences at the time, but which he didn't gather.

Chapter 8

Conclusion

This brings us to the end of our exploration of the value of information and its relationship to the epistemology of individual inquiry. What, then, is the view? As noted at the end of the previous chapter, the Value of Information framework that grows out of Hosiasson's initial insight centres on a way of evaluating and comparing different evidence-gathering episodes with one another, with the choice not to gather evidence, and with other possible courses of action; and it provides a way of doing this evaluation from a pragmatic point of view, an epistemic point of view, and a combination of the two, which I've been calling the all-things-considered point of view.

The ingredients of this evaluative framework are as follows:

(I) *the vantage point from which the evaluation takes place*.

This includes:

(a) *a probability function*.

Typically, this is the would-be evidence-gatherer's credences when they must decide what to do; but equally it could be the evidential probabilities, if such exist, or the credences the would-be evidence-gatherer would have had had they behaved fully rationally in the past.

(b) *either*

(i) *a set of decision problems and a pragmatic utility function*.
The decision problems are those the individual thinks they might face with their credences, and the pragmatic utility function encodes their values and ends. These ingredients are needed for the pragmatic evaluation of evidence-gathering episodes.

or

(ii) *an epistemic utility function*.
This is needed for the epistemic evaluation of evidence-gathering episodes.

(II) *the available options*.

These might include evidence-gathering episodes represented by evidence functions, including the trivial evidence function, which represents not gathering evidence at all, but they might also include non-epistemic actions, such as making a sandwich.

(III) *the updating plan*.

This describes how the individual will respond to the evidence.

The crucial insight of the approach that descends from Hosiasson's original insight is that an evidence-gathering episode is an action like any other. She, Blackwell, and Good saw how to assess it for its pragmatic value, and Oddie saw how to evaluate it for its pure epistemic value. This allows us to bring such decisions within the ambit of the totalizing vision of rational choice theory, which seeks to govern at least the rationality of all of our actions.

The norms of evidence-gathering and its extended pursuit, which we call inquiry, are then simply instances of the more general norms of rational choice. If the correct theory of decision is expected utility theory, then these are:

Pragmatic norm of inquiry Gather evidence when doing so maximizes expected pragmatic utility.

Epistemic norm of inquiry Gather evidence when doing so maximizes expected epistemic utility.[1]

All-things-considered norm of inquiry Gather evidence when doing so maximizes expected all-things-considered utility.

But the framework is flexible enough that it will serve if a different theory of decision is correct, such as one that permits sensitivity to risk or ambiguity or both (recall Section 2.5).

What's more, by expanding our understanding of the possibilities over which credences and utilities are defined in our decision-making model so that they include impossibilities as well as possibilities, and by representing logical or *a priori* reasoning as ruling out such impossibilities, just as empirical evidence rules out empirical possibilities, we can understand logical and *a priori* inquiry using the Value of Information framework, and thereby provide a unified account of all inquiry, empirical and otherwise.

[1] Flores & Woodard (2023) ask whether there are genuinely epistemic norms on evidence-gathering. This seems a candidate for this. In the end, I don't think too much hangs on how we categorize norms into the pragmatic and the epistemic. The norm just given is epistemic in the sense that it evaluates something for how it serves purely epistemic ends. But the thing evaluated is not itself purely epistemic: it is not a credence or a belief, but rather an action. The action is of course aimed at gathering evidence, which in turn is aimed at changing a purely epistemic state, such as a credence or belief. Is that sufficient to make it epistemic? As I say, I'm not sure we need adjudicate this. The insights of Flores and Woodard's paper, which describes epistemic ways in which we criticize evidence-gathering episodes that violate certain norms, stand whether or not they support the claim that the norms themselves are epistemic.

Bibliography

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica*, *21*(4), 503–546.

Archer, A. (13). The Aim of Inquiry. *Disputatio*, *61*(95-119).

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York: Springer.

Blackwell, D. (1951). Comparison of Experiments. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, *2*, 93–102.

Bottomley, C., & Williamson, T. L. (ta). Rational risk-aversion: Good things come to those who weight. *Philosophy and Phenomenological Research*.

Bright, L. K. (2017). On Fraud. *Philosophical Studies*, *174*(2), 291–310.

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

Buchak, L. (2013). *Risk and Rationality*. Oxford, UK: Oxford University Press.

Carr, J. (2021). A Modesty Proposal. *Synthese*, *198*(4), 3581–3601.

Das, N. (2023). The Value of Biased Information. *The British Journal for the Philosophy of Science*, *74*(1), 25–55.

Dorst, K. (2020). Evidence: A Guide for Uncertain. *Philosophy and Phenomenological Research*, *100*(3), 489–757.

Dorst, K., Levinstein, B. A., Salow, B., Husic, B. E., & Fitelson, B. (2021). Deference Done Better. *Philosophical Perspectives*, *35*, 99–150.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics*, *75*(4), 643–69.

Falbo, A. (2023). Should epistemology take the zetetic turn? *Philosophical Studies*, *180*(10-11), 2977–3002.

Flores, C., & Woodard, E. (2023). Epistemic Norms on Evidence-Gathering. *Philosophical Studies*, *180*, 2547–2571.

Friedman, J. (2020). The Epistemic and the Zetetic. *Philosophical Review*, *129*(4), 501–536.

Gallow, J. D. (2021). Updating for Externalists. *Noûs*, *55*(3), 487–516.

Gardiner, G. (2022). Attunement: On the Cognitive Virtues of Attention. In M. Alfano, C. Klein, & J. de Ridder (Eds.) *Social Virtue Epistemology*. New York: Routledge.

Geanakoplos, J. (1989). Game Theory without Partitions, and Applications to Speculation and Consensus. Tech. rep., Cowles Foundation Discussion Paper.

Gibbard, A., & Harper, W. (1978). Counterfactuals and Two Kinds of Expected Utility. In E. F. M. Clifford Alan Hooker, James L. Leach (Ed.) *Foundations and Applications of Decision Theory*, vol. 13a of *University of Western Ontario Series in Philosophy of Science*, (pp. 125–162). D. Reidel.

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.

Hacking, I. (1967). Slightly more realistic personal probability. *Philosophy of Science*, *34*(4), 311–325.

Hookway, C. (2006). Epistemology and Inquiry: The Primacy of Practice. In S. Hetherington (Ed.) *Epistemology Futures*, (pp. 95–110). Oxford University Press.

Hosiasson, J. (1931). Why do we prefer probabilities relative to many data? *Mind*, *40*(157), 23–36.

Huttegger, S. M. (2013). In Defense of Reflection. *Philosophy of Science*, *80*(3), 413–433.

Huttegger, S. M. (2014). Learning experiences and the value of knowledge. *Philosophical Studies*, *171*(2), 279–288.

Huttegger, S. M., & Nielsen, M. (2020). Generalizing Learning and Conditional Expectation. *Philosophy of Science*, *87*(5), 868–883.

Isaacs, Y., & Russell, J. S. (2023). Updating without Evidence. *Noûs*, *57*(3), 576–599.

Jeffrey, R. C. (1983). *The Logic of Decision*. Chicago and London: University of Chicago Press, 2nd ed.

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*. Dordrecht and Heidelberg: Springer.

Kelp, C. (2021). *Inquiry, Knowledge, and Understanding*. Oxford, UK: Oxford University Press.

Levinstein, B. A. (2023). Accuracy, Deference, and Chance. *Philosophical Review*, *132*(1), 43–87.

Lewis, D. (1981). Causal Decision Theory. *Australasian Journal of Philosophy*, *59*, 5–30.

Mayo-Wilson, C., Zollman, K. J. S., & Danks, D. (2011). The Independence Thesis: When Individual and Social Epistemology Diverge. *Philosophy of Science*, *78*(4), 653–77.

Myrvold, W. C. (2012). Epistemic values and the value of learning. *Synthese*, *187*(2), 547–568.

Nielsen, M. (2024). Only CDT values knowledge. *Analysis*, *84*(1), 67–82.

O'Connor, C., & Weatherall, J. O. (2018). *The Misinformation Age: How False Beliefs Spread*. New Haven: Yale University Press.

Oddie, G. (1997). Conditionalization, Cogency, and Cognitive Value. *British Journal for the Philosophy of Science*, *48*, 533–41.

Pérez Carballo, A. (2018). Good Questions. In J. Dunn, & K. Ahlstrom-Vij (Eds.) *Epistemic Consequentialism*. Oxford, UK: Oxford University Press.

Pettigrew, R. (2020). Logical Ignorance and Logical Learning. *Synthese*.

Pettigrew, R. (2023a). Bayesian updating when what you learn might be false. *Erkenntnis*, *88*, 309–324.

Pettigrew, R. (2023b). Epistemic Utility Arguments for Epistemic Norms. In E. N. Zalta (Ed.) *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Quiggin, J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization*, *3*, 323–43.

Quiggin, J. (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*. Kluwer Academic Publishers.

Ramsey, F. P. (1990). Weight or the Value of Knowledge. *The British Journal for the Philosophy of Science*, *41*, 1–4.

Rosa, L. (2025). *Rules for the Inquiring Mind: A Unified Framework of Norms of Inquiry*. London, UK: Routledge.

Rosenstock, S., Bruner, J., & O'Connor, C. (2017). In Epistemic Networks, Is Less Really More? *Philosophy of Science*, *64*(2), 234–252.

Salow, B. (2018). The externalist's guide to fishing for compliments. *Mind*, *127*, 691–728.

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

Schliesser, E. (ta). Synthetic Philosophy: a Restatement. *Proceedings of the Aristotelian Society*.

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

Schultheis, G. (ta). Accurate Updating. *Philosophy of Science*.

Simion, M. (2023). Resistance to Evidence and the Duty to Believe. *Philosophy and Phenomenological Research*.

Skyrms, B. (1990a). *The Dynamics of Rational Deliberation*. Cambridge, Mass.: Harvard University Press.

Skyrms, B. (1990b). The Value of Knowledge. *Minnesota Studies in the Philosophy of Science*, *14*, 245–66.

Staffel, J. (2019). Credences and suspended judgments as transitional attitudes. *Philosophical Issues*, *29*(1), 281–294.

Staffel, J. (2021a). Pro tem rationality. *Philosophical Perspectives*, *35*(1), 383–403.

Staffel, J. (2021b). Transitional attitudes and the unmooring view of higher-order evidence. *Noûs*, *57*(1), 238–260.

Staffel, J. (ta). *Unfinished Business: Rational Attitudes in Reasoning*. Oxford, UK: Oxford University Press.

Thorstad, D. (2022). There are no epistemic norms of inquiry. *Synthese*, *200*(410).

Torsell, C. (2024). Janina Hosiasson and the value of evidence. *Studies in the History and Philosophy of Science*, *106*, 31–36.

van Fraassen, B. C. (1984). Belief and the Will. *Journal of Philosophy*, *81*, 235–56.

van Fraassen, B. C. (1995). Belief and the Problem of Ulysses and the Sirens. *Philosophical Studies*, *77*(1), 7–37.

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

Willard-Kyle, C. (forthcoming). The Knowledge Norm of Inquiry. *The Journal of Philosophy*.

Williams, J. R. G. (2018). Rational Illogicality. *Australasian Journal of Philosophy*, *96*(1), 127–141.

Williams, J. R. G., & Pettigrew, R. (2023). Consequences of Calibration. *The British Journal for the Philosophy of Science*.

Williamson, T. (2013). Gettier Cases in Epistemic Logic. *Inquiry: An Interdisciplinary Journal of Philosophy*, *56*(1), 1–14.

Williamson, T. (2014). Very Improbably Knowing. *Erkenntnis*, *79*(5), 971–999.

Woodard, E. (2022). Why Double-Check? *Episteme*, (pp. 1–24).

Zollman, K. J. S. (2007). The communication structure of epistemic communities. *Philosophy of Science*, *74*(5), 574–587.

Zollman, K. J. S. (2010). The Epistemic Benefit of Transient Diversity. *Erkenntnis*, *72*, 17–35.