Rational Aversion to Information

Sven Neth

Forthcoming in *The British Journal for the Philosophy of Science**

September 2023

Abstract

Is more information always better? Or are there some situations in which more information can make us worse off? Good (1967) argues that expected utility maximizers should always accept more information if the information is cost-free and relevant. But Good's argument presupposes that you are certain you will update by conditionalization. If we relax this assumption and allow agents to be uncertain about updating, these agents can be rationally required to reject free and relevant information. Since there are good reasons to be uncertain about updating, rationality can require you to prefer ignorance.

1 Introduction

We care about learning the truth for its own sake, but we also care about learning because it can lead us to make better decisions. That is, besides the epistemic benefits of finding out the truth, learning often comes with instrumental benefits as well.

Is more information always instrumentally better? Or are there situations in which more information can make us foreseeably worse off? It is clear that information can make us worse off if we consider the cost of processing and

---

storing the information or the opportunity cost of thinking for too long before acting. Nobody thinks that you have to read all the reviews before buying a new vacuum cleaner or that you should think long before hitting the brakes when a red light comes up. It is also clear that information can make us worse off if it is false, so let me be clear that when I talk about information, I always mean true information.

What if the information is cost-free? For rational agents, is it always instrumentally valuable to accept free information? Good (1967) argues that the answer is 'yes' if we accept the principle of maximizing expected utility. However, Good presupposes that you are certain you will update by conditionalization, which means you are certain your new credences after learning are equal to your old conditional credences given the learned event. There are good reasons to assign positive probability to failures of conditionalization, even for rational agents. I show that if you assign a positive probability to failures of conditionalization, the principle of maximizing expected utility can require you to reject free information. Sometimes, even expected utility maximizers are better off knowing less. Moreover, this offers a vindicating explanation of why people sometimes reject information in real-life examples, such as medical testing.

To be clear, this paper is not about situations in which you actually fail to conditionalize. In all my examples below, we can assume that the agent conditionalizes in the actual world. Rather, this paper is about situations in which you fail to be certain that you will conditionalize. You can fail to be certain that you will conditionalize even if you always conditionalize.

Here is the plan. First, I explain Good's argument. Then, I explain why Good's argument presupposes that you are certain you will update by conditionalization and give reasons to reject this assumption. I show how assigning a positive probability to failures of conditionalization can make it rational to reject free information for expected utility maximizers and sketch how this can explain information aversion in the real world. I finish by explaining how we can generalize the value of information to agents who are uncertain about how they will update.

2

2 Good's Argument

I start by introducing some terminology and explain Good's argument.

2.1 Terminology

I use the framework of Savage (1972) to model decision making under uncertainty. We have a set Ω of states, which contains all epistemically possible worlds from the point of view of the agent we are modeling. Events are subsets of Ω and we model the credences of our agent by a probability function.[1] We also have a set $\mathscr{O}$ of outcomes, where outcomes contain everything our agent cares about. We model our agent's preferences over outcomes by a utility function u which maps outcomes to their utilities. Actions are functions from states to outcomes. I assume actions are causally and probabilistically independent of states.[2]

Given a probability function p and utility function u, the expected utility of action f is:

$$\mathbb{E}_p(f) = \sum_{\omega \in \Omega} p(\omega) u(f(\omega)).^3$$

I assume your utility function remains fixed through learning.[4] However, your credences change in response to evidence, so I relativize expected utility to a probability function.

A choice set is a set of actions among which our agent makes a decision. I assume that all choice sets are finite. Our agent maximizes expected utility if for every available choice set $\mathcal{S} = \{f_1, ..., f_n\}$, she picks an action $f_i \in \mathcal{S}$ which maximizes expected utility relative to her probability function p and utility function u.

I model learning by an evidence partition $\mathcal{E}$ of Ω. This partition contains

---

[1] I assume Ω is finite and model credences as finitely additive probability function $p : \mathcal{P}(\Omega) \to [0,1]$.

[2] Adams and Rosenkrantz (1980) and Maher (1990) discuss how Good's argument fails if this assumption is relaxed, in both evidential and causal decision theory.

[3] $p(\omega)$ is shorthand for $p(\{\omega\})$.

[4] I set aside cases in which learning leads you to change your utility function, perhaps in a 'transformative experience' (Paul 2014; Pettigrew 2019).

the events our agent might learn, which are mutually exclusive and collectively exhaustive. We can think of the evidence partition as a question, for example the question whether it is sunny or rainy outside. In this case, the evidence partition contains two cells: the worlds where it is sunny outside and the worlds where it is rainy outside. Since the events in the evidence partition are live possibilities for what our agent might learn, they all have non-zero probability, so $p(E) > 0$ for all $E \in \mathcal{E}$.

When learning event E in the evidence partition, our agent updates her credences to $\mathcal{P}_E$. I allow our agent to be uncertain about how she will update. This means that $\mathcal{P}_E$ is not a particular probability function, but rather a random variable whose values can be different probability functions. (Hence the fancy typeface.) The only constraint I impose is that after learning an event, our agent is certain of that event.[5] I write $p(\cdot \mid E)$ for the credences our agent adopts after learning event $E \in \mathcal{E}$ and updating by conditionalization.[6]

Here is an example of our framework in action. You are at the horse track thinking about which horse to bet on. The states specify which horse will win the race (and other relevant facts) and the actions are different bets you might place. The probability function p encodes your credences about different horses winning and the utility function u models how much you value the outcomes of these bets, for example different amounts of money.

Imagine a charming stranger comes up to you and offers you their opinion on which horse is likely to win. Are you willing to listen? The evidence partition $\mathcal{E}$ contains different opinions the stranger might voice and $\mathcal{P}_E$ models how you expect to update your credences after listening. To be clear, the information you learn is not which horse is likely to win but only what the stranger is saying. The stranger might be lying or clueless.

You need to decide: Do you want to find out what the stranger has to say or would you rather place your bet now? It is not obvious how to answer this question. On the one hand, you might listen to the stranger and ignore

---

[5]Let $\Delta(\Omega)$ be the set of all probability functions $p : \mathcal{P}(\Omega) \to [0, 1]$. Formally, $\mathcal{P}_E$ is a function from E to $\Delta(\Omega)$ such that for each $\omega \in E$, $\mathcal{P}_E(\omega)(E) = 1$. For each $\omega \in E$, $\mathcal{P}_E(\omega)$ is a particular probability function.

[6]I use the standard ratio definition: $p(A \mid E) = \frac{P(A \cap E)}{p(E)}$ assuming $p(E) > 0$.

what they say if you do not find it helpful, so how could listening harm you? On the other hand, perhaps the stranger is trying to mislead you. In this case, do you trust yourself to listen before placing your bet?

2.2 The Argument

Good (1967) thinks you should listen to the stranger before placing your bet. More generally, Good argues that if you are rational, then given any choice set and evidence partition, you are never worse off by first learning the true event in the evidence partition and making your choice afterwards rather than making your choice now.[7] Good does not mean that more information always leads to better decisions. You might get unlucky and learn something misleading. Rather, Good argues that learning cannot make you foreseeably worse off.

The idea behind Good's argument can be illustrated by an example. Suppose there is a race between horse A and horse B tomorrow. You have to decide whether (i) to bet on horse A, which means you win \$1 if A wins and lose \$2 otherwise, (ii) to bet on horse B, which means you win \$1 if B wins and lose \$2 otherwise, or (iii) to play it safe, which means you won't win or lose anything. You think A and B are equally likely to win, so your best option right now is to play it safe. But you can listen to the (accurate) weather report for tomorrow. You think that A is $\frac{3}{4}$ likely to win if it rains and B is $\frac{3}{4}$ likely to win if the sun shines. We can illustrate your decision problem as shown in figure 1, where rectangles stand for decisions you face ('choice nodes') and circles stand for events which might happen ('chance nodes').

Here is a more general explanation. Good assumes that rational agents maximize expected utility. So, if you are rational, then given some choice set $\mathcal{S}$, you will choose what seems best by your current lights: an action in $\mathcal{S}$

---

[7]Skyrms (1990) provides a helpful overview and points out that Ramsey (1990) and Savage (1972) give similar arguments. Good (1967) notes that his argument is partly anticipated by Raiffa and Schlaifer (1961, p. 90) and Lindley (1965, p. 66). Hosiasson (1931) discusses similar ideas and cites an unpublished paper by Ramsey as inspiration. Interestingly, the argument does not work when you decide whether someone else should learn more information before making a decision (Good 1974).
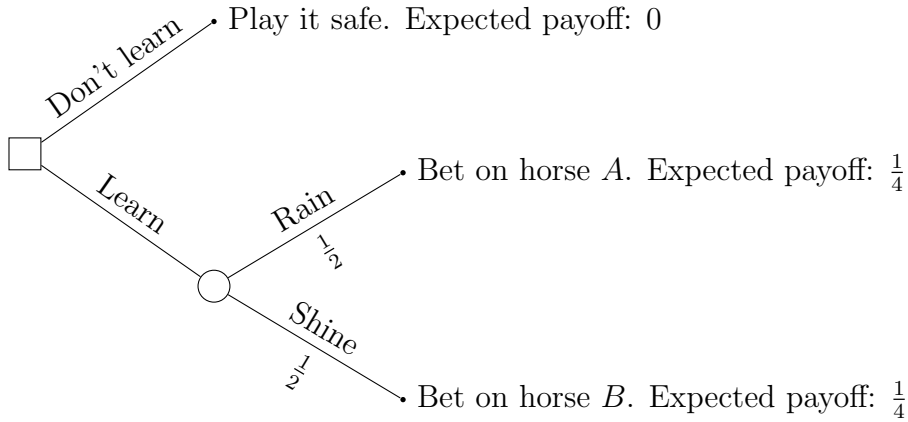
Figure 1: If you care about winning, you should listen to the weather report.

which maximizes expected utility with respect to your current credences. So the expected value of choosing now is the expected utility of one of the best actions in $\mathcal{S}$ relative to your current credences p:

$$\max_{f \in \mathcal{S}} \mathbb{E}_p(f).$$

If, on the other hand, you learn that E is the true event in our evidence partition, you update your credences p to $\mathcal{P}_E$. Good assumes that in each state with non-zero probability, your updated credences are obtained from your current credences by conditionalizing, so $\mathcal{P}_E = p(\cdot \mid E)$. Good also assumes that learning is cost-free. This means that before and after learning, you choose among the same actions and outcomes have the same utilities. The only impact of the information is to change your credences.[8]

After learning, you choose what seems best by your updated lights—an action in $\mathcal{S}$ which maximizes expected utility with respect to your updated credences:

$$\max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot \mid E)}(f).$$

[8]Consider cases in which the information is not free (processing costs, library fees). In such cases, outcomes before and after learning do not have the same utility. Kadane, Schervish, and Seidenfeld (2008, pp. 17-20) discuss this issue in detail. You might also ascribe negative utility to the information itself, for example because it makes you feel bad (Golman, Hagmann, and Loewenstein 2017). I set such cases aside and focus on the instrumental value of information.

You don't know yet which element of the evidence partition you will learn. But we can consider the expected value of acting after learning:

$$\sum_{E \in \mathcal{E}} p(E) \max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot|E)}(f).$$

Good completes the argument by proving that the expected value of acting after learning is always greater than or equal to the expected value of choosing now:

$$\sum_{E \in \mathcal{E}} p(E) \max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot|E)}(f) \geq \max_{f \in \mathcal{S}} \mathbb{E}_p(f).$$

Moreover, this inequality is strict unless there is some action $f \in \mathcal{S}$ which is best regardless of which event in the evidence partition you learn—that is, unless the evidence partition is irrelevant for the choice set under consideration. So according to Good, the principle of maximizing expected utility entails:

Value of Learning:

    i. Rational agents are always permitted to accept free information before making a decision.

    ii. Rational agents are always required to accept free and relevant information before making a decision.

2.3 What does the Argument show?

Does Good's argument show that Value of Learning is correct? There are ways to push back. One might question the assumption that rational agents always maximize expected utility (Buchak 2010; Campbell-Moore and Salow 2020) or that rationality requires precise credences (Bradley and Steele 2016; Wheeler 2021).[9] Relaxing these assumptions leads to cases where you can be

---

[9]One could also question the assumption that learning can always be modeled as learning an element of a partition (Salow and Ahmed 2019; Dorst 2020; Das 2023), which is foreshadowed by Williamson (2000, pp. 230-7). Merely finitely additive probabilities can also lead to information aversion (Kadane, Schervish, and Seidenfeld 1996, 2008). Since I restrict attention to finite state spaces, the debate over finite versus countable additivity does not concern me here.

required to reject free information. One might take this to question Value of Learning. However, one might also take this as a strike against alternatives to expected utility theory with precise credences.

One might think that the permissibility of accepting free information before making a decision is independently plausible, a piece of common sense: 'look before you leap'. It is a mark in favor of expected utility theory that it entails this piece of common sense and a problem for other decision theories if they conflict with it. From this perspective, Good's argument is not really an argument for Value of Learning but rather an argument for expected utility theory. This interpretation is suggested by Kadane, Schervish, and Seidenfeld (2008):

> So the question remains of whether it is reasonable to impose the requirement on a theory of rational decision making that it not require or permit paying not to see cost-free data. If it is, the only such theory known to us is Bayesian decision theory with a single countably-additive proper prior. (Kadane, Schervish, and Seidenfeld 2008, p. 33)

Arguments along these lines are common. The general shape of the argument is that (a) Value of Learning is correct and (b) alternatives to expected utility theory are bad because they conflict with this. This assumes that (c) expected utility theory entails Value of Learning.[10]

These arguments are misguided. At least, they require serious qualification. This is because expected utility theory, supplemented with plausible assumptions, entails that there are cases in which we are rationally required to reject free information. Good's argument rests on the auxiliary assumption

---

[10]More examples: Wakker (1988) shows that violating the independence axiom of expected utility theory leads to situations in which agents reject free information and takes this to show that such violations are irrational. Al-Najjar and Weinstein (2009, p. 249) object to decision theories allowing for ambiguity aversion because they rationalize aversion to information "which most economists would consider absurd or irrational". Briggs (2015) and Ahmed (2016) object to risk-weighted expected utility (REU) theory because it leads to diachronic inconsistency and aversion to information. Buchak (2013, pp. 187–9) also discusses how the value of information can be negative in REU theory and considers this to be a serious cost.

that you are certain you will update by conditionalization. This assumption should not be built into expected utility theory and there are good reasons to reject it. If we reject this assumption, expected utility maximizers with precise credences can be required to reject free and relevant information. So (c) is false: expected utility theory does not entail Value of Learning. I also argue that (a) is false: rational agents can be required to reject free information. So we should not take Value of Learning as axiomatic in our theories of instrumental rationality.

If you are already skeptical of Value of Learning, you might argue as follows: expected utility theory entails Value of Learning but Value of Learning is clearly false. There are many situations in real life where we are better off ignoring free information. Perhaps you think that the stranger at the horse track will try to deceive you. Therefore, we should reject expected utility theory and look for an alternative decision-theoretic framework, perhaps risk-weighted expected utility theory or imprecise credences. I agree that there are many situations in real life where we are better off ignoring free information. However, once we understand that expected utility theory does not entail Value of Learning, we can make sense of information aversion within the standard framework of expected utility theory and Bayesian epistemology.

3   Against Good's Argument

I explain why Good's argument presupposes that you are certain you will conditionalize (Immodesty). I argue that this assumption is implausible. Instead, we should assign some positive probability to not conditionalizing (Modesty). I show that expected utility maximization can require modest agents to reject free information.

3.1   Good presupposes Immodesty

As we have seen, Good's argument requires that whichever event in the evidence partition you learn, your future credences are obtained from your

current credences by conditionalization:

The Equation: $\mathcal{P}_E = p(\cdot \mid E)$ for every $E \in \mathcal{E}$.[11]

At first glance, one might think The Equation means that you are actually a conditionalizer. This is how the assumption is sometimes glossed in presentations of Good's argument.[12] However, The Equation actually means that you assign subjective probability one to the event that you will conditionalize.[13] In other words, you are certain you will conditionalize. This is because The Equation says that for every event you might learn, your new credences equal your old credences conditional on the learned event. Taken together, the events in the evidence partition sum to probability one. This means that in every state with positive probability, your new credences equal your old credences conditional on the true event in the evidence partition. Since states represent epistemic possibilities, you are certain you will conditionalize.

So Good's argument presupposes

Immodesty: You are certain you will conditionalize.

Here is another way to bring this out. You might in fact update by conditionalization. Nonetheless, you might assign positive probability to a state in which you fail to conditionalize. In this case, The Equation does not hold and Good's argument does not go through. On the other hand, you might be certain you will conditionalize—and so satisfy The Equation— but fail to conditionalize in the actual world. This might be because you have assigned probability zero to an unforeseen failure of rationality. In this case, Good's argument still applies. What is at issue is not whether you will actually conditionalize but whether you are certain you will conditionalize. Even if you always conditionalize, you might have good reasons not to be certain of that.

---

[11]Technically, Good's result requires only that this equality holds with probability one.

[12]For example, Laffont (1989, p. 58) presents a result equivalent to Good's and writes that the agent under consideration "revises his expectations by using Bayes's theorem". This sounds like we're assuming that the agent is actually a conditionalizer.

[13]As Skyrms (1990, p. 247) writes, "the proof implicitly assumes not only that the decision maker is a Bayesian but also that he knows he will act as one. The decision maker believes with probability one that if he performs the experiment he will [...] update by conditionalization [...]". Huttegger (2014, p. 283) also makes this point.

3.2 The Case for Modesty

Immodesty is implausible. Instead, we should accept:

> **Modesty:** There is some positive probability that you will not
> conditionalize.

To be clear, what I have in mind here is subjective probability, not objective
chance. So to accept Modesty means to assign some positive credence to the
possibility that you will not conditionalize.[14]

There are good reasons for Modesty. Moreover, these reasons flow from
standard principles of Bayesian epistemology. Let me first be clear that it
is by no means (physically or metaphysically) necessary that you will condi-
tionalize. Rather, the claim that your new credences after learning are your
old credences conditional on the learned event is a substantive claim about
how your credences will evolve over time. The following passage by Ramsey
makes the point clear:

> [the degree of belief in p given q] is not the same as the degree to
> which [a subject] would believe p, if he believed q for certain; for
> knowledge of q might for psychological reasons profoundly alter
> his whole system of beliefs. (Ramsey 1931, p. 21)[15]

So it is a consistent possibility that you fail to conditionalize. Many Bayesians
are attracted to the principle of regularity, which says that you should assign

---

[14]Modesty has been defended before. For example, discussing whether we should defer
to our future credences, Briggs (2009, pp. 59–60) writes: "Under all but the most ideal
circumstances, agents will have reasons to suspect that future failures of conditionalization
are in store". Pettigrew (2020) points out that standard arguments for conditionalization
assume 'deterministic updating' and so leave no room for uncertainty about how you will
update. Lederman (2015) draws on failures of common knowledge that we will conditional-
ize to construct counterexamples to Aumann's claim that rational agents cannot 'agree to
disagree'. Cohen (2020) discusses uncertainty about updating in the context of epistemic
logic. Christensen (2007, p. 3) defends the broader claim that "even an agent who is in
fact cognitively perfect might, it would seem, be uncertain of this fact". Similar ideas are
defended by many others, including Carr (2019), Bradley (2020), and Dorst (2020).

[15]Diaconis and Zabell (1982) discuss this passage. Of course, the general idea is much
more broadly recognized. For example, in *The Portrait of a Lady*, Henry James writes
about some piece of news: "But it had been one thing to foresee it mentally, and it was
another to behold it actually" (James [1882] 2011, p. 217).

positive prior probability to all consistent possibilities. This principle entails Modesty.

More broadly, that you will conditionalize is an empirical proposition. Rationality should not require you to be certain that some empirical proposition is true. For example, if you suffer brain damage as the result of a stroke, you will likely not conditionalize. Plausibly, you should not be certain that you won't suffer brain damage in the future. Therefore, you should not be certain that you will conditionalize.[16]

We can make an even stronger case for Modesty. There is a long research tradition in psychology and cognitive science which aims to demonstrate that humans are not ideal Bayesian agents and so do not always conditionalize. There are a number of well-documented fallacies and heuristics which deviate from conditionalization. An example is the base rate fallacy, in which people ignore prior probabilities and so overestimate the probability of rare events (Kahneman and Tversky 1973). Another example is the gambler's fallacy, which is when people think that a fair coin landing heads provides evidence that the next coin flip will land tails.

Once you learn about these empirical findings, it seems reasonable to believe that they might also apply to yourself. Therefore, you should assign some positive credence to not conditionalizing and accept Modesty. In addition to such general considerations, you might remember specific cases in which you did not conditionalize, but committed (say) the gambler's fallacy. If you have such evidence, this gives you another strong reason for Modesty.

Perhaps you are quite confident of your future rationality. But even if you currently have no evidence that you might fail to conditionalize, it seems reasonable that you might obtain such evidence. For example, you might learn that you just took a drug which increases your susceptibility to the gambler's fallacy or that your brain is wired to misfire in certain situations.[17] Surely, if you learned something like this, you should decrease your credence that your future self will conditionalize. But Immodesty rules

---

[16]Note that the possibility of malfunction does not only apply to humans, but also to AI agents and plausibly to any kind of agent which is physically realized.

[17]The reason-impairing drug is inspired by Christensen (2007).

this out: once you assign probability zero to failures of conditionalization, then no matter what you learn, you will continue to assign probability zero to failures of conditionalization (if you actually conditionalize).[18] This seems unreasonable—surely, there are some things you might learn that would make you doubt your own future rationality. Therefore, you should assign non-zero probability to failures of conditionalization, so Modesty follows.

The arguments above appeal to substantive constraints on prior probabilities. Subjective Bayesians reject such constraints beyond adherence to the axioms of probability. So subjective Bayesians will not be moved by my arguments. However, Immodesty is also a substantive constraint on prior probabilities and does not follow from the axioms of probability. Subjective Bayesians have no reason to accept this constraint.[19]

There are, of course, good reasons to think that rationality requires conditionalization, for example diachronic coherence arguments (Lewis 1999) and various accuracy arguments (Joyce 1998; Greaves and Wallace 2005; Pettigrew 2016). Modesty is entirely consistent with this claim. Arguments for conditionalization aim to show:

Conditionalization: You should conditionalize.

Good's argument relies on Immodesty, which says that you are certain you will conditionalize. Conditionalization does not entail Immodesty. We can accept that we should conditionalize but still have good reason to assign positive probability to failures of conditionalization in the future. This is because we might not be certain that our future selves will be rational. Indeed, as good Bayesians, we should not be certain that our future selves will be rational if our evidence suggests that we might not be.

---

[18]This is the key reason Bayesian epistemologists tend to be skeptical of assigning probability zero to any possible event. For example, Lewis (1980, p. 268) argues that regularity is "required as a condition of reasonableness: one who started out with an irregular credence function (and who then learned from experience by conditionalizing) would stubbornly refuse to believe some propositions no matter what the evidence in their favor."

[19]As Hacking (1967, p. 315) points out, the axioms of probability don't entail that you will actually conditionalize, much less that you are certain of doing so: "The idea of the model of learning is that $Prob(h/e)$ represents one's personal probability after one learns e. But formally, the conditional probability represents no such thing. [...] $Prob(h/e)$ stands merely for the quotient of two probabilities."

Some philosophers have argued that the arguments for conditionalization only support the norm that you should intend or plan to conditionalize, rather than the norm that you should actually conditionalize.[20] If these philosophers are correct, then it is even harder to see any conflict between the arguments for conditionalization and Modesty. We can rationally plan to ϕ while also thinking that there is some positive probability that we will fail to ϕ. For example, I can plan to run a 10K race while also thinking that there is some chance I won't make it to the end.[21]

There are also reasons to doubt whether conditionalization is always rationally required. For example, Douven (2013) argues that an alternative to conditionalization, which he calls 'Inference to the Best Explanation', leads you to converge to the truth faster in some circumstances. If you care about fast convergence, this might be a reason to use Douven's 'Inference to the Best Explanation' instead of conditionalization. While this is no conclusive argument against conditionalization, it might perhaps instill some doubt about whether conditionalization is always rational. Plausibly, the right response to this normative uncertainty is to assign some positive probability to failures of conditionalization even if you are sure you will update rationally.

3.3   Modesty entails Information Aversion

Let us assume Modesty. I now explain how for modest agents, maximizing expected utility can require you to reject free information. The basic idea is quite simple. If you are modest, you assign some credence to the possibility that learning more information will lead you to make inferences which you do not currently endorse. This might lead you to make choices which, from your current point of view, seem like a bad idea. Therefore, you might be better off ignorant.

Suppose a fair coin will be flipped twice and Ann knows this. She chooses among bets on the second coin flip: a safe bet which always yields zero, a

---

[20]This point is discussed, for example, by Pettigrew (2016, pp. 187-88).
[21]Bratman (1992, p. 11-12) discusses similar examples and argues that one can plan to ϕ without believing that one will ϕ.

risky bet on heads and a risky bet on tails. Our choice set $\mathcal{S}$ is:

safe : {\$0 always},

risky-heads : {\$1 if the second coin flip lands heads, $-\$2$ otherwise},

risky-tails : {\$1 if the second coin flip lands tails, $-\$2$ otherwise}.

Ann values money linearly and is an expected utility maximizer. She is also certain that her future self will be an expected utility maximizer.[22]

I offer Ann the following choice: She can either make her decision now or learn the outcome of the first coin flip and make her decision afterwards. If Ann makes her decision now, she will pick safe. So the expected value of choosing now is:

$$\max_{f \in \mathcal{S}} \mathbb{E}_p(f) = \mathbb{E}_p(\texttt{safe}) = 0.$$

What happens if Ann learns the outcome of the first flip and makes her decision afterwards? If Ann conditionalizes, she will choose the safe bet no matter what she learns since she regards the two coin flips as independent. So there is no reason for her to avoid learning. It can't help her, but it can't hurt her either.

But Ann is modest and assigns some positive probability to failures of conditionalization. In particular, Ann assigns some positive probability to committing the gambler's fallacy: after she learns that the first coin flip lands heads, she will become confident that the next coin flip will land tails and vice versa. In particular, Ann assigns some positive probability ϵ to the event that when she learns that the first coin flip lands heads, she will become .9

---

[22]I mean that she is certain she will pick one of the actions in $\mathcal{S}$ which maximizes expected utility relative to her updated credences—which might or might not be obtained from her current credences by conditionalization. The value Ann currently assigns to f on the supposition of E is the conditional expected utility $\mathbb{E}_{p(\cdot|E)}(f) = \sum_{\omega \in \Omega} p(\{\omega\} \mid E)f(\omega)$. Gyenis and Rédei (2017) discuss conditional expectations in a much more general setting. The important point is that this conditional expected utility can come apart from the value Ann assigns to f after actually learning E. I also assume that $\mathcal{S}$ does not include actions like 'adopt credence p after learning evidence E'. With such an extended option set, one can argue that certainty that one will maximize expected utility entails certainty that one will conditionalize (Brown 1976), although Pettigrew (2020) points out how uncertainty about updating complicates this argument. Thanks to an anonymous referee for pushing me to clarify how exactly I understand certainty that one maximizes expected utility.

confident that the second coin flip will land tails and vice versa.

Now suppose Ann learns that the first coin flip lands heads and commits the gambler's fallacy. Relative to her updated credences, the risky bet on tails now looks like the best option. However, given Ann's current credences, the risky bet is the wrong choice. The situation is analogous if Ann learns that the first coin flip lands tails and commits the gambler's fallacy. Figure 2 sums up Ann's situation.
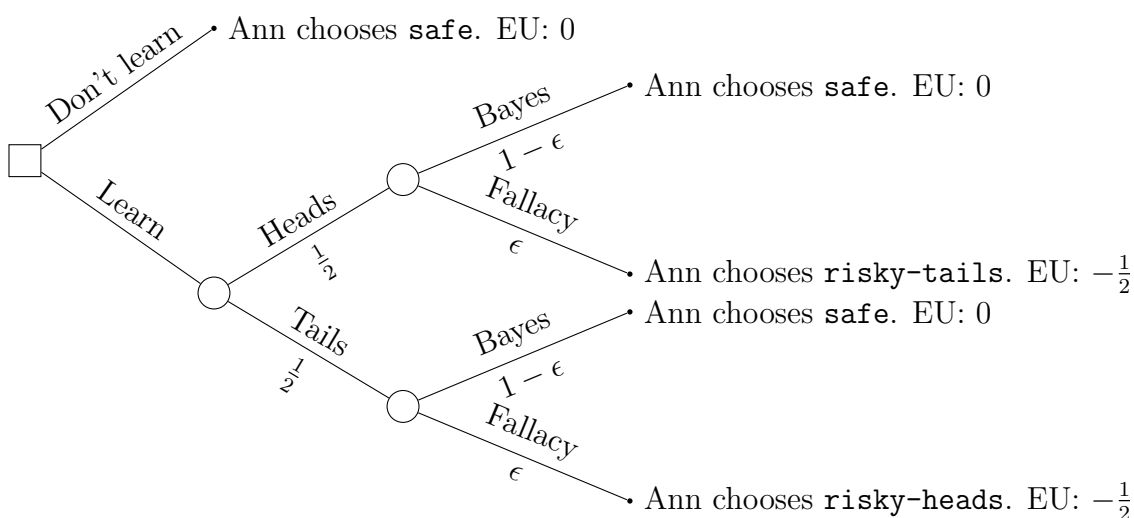

Figure 2: Ann's decision problem.

The expected value of learning and deciding afterwards is $-\frac{\epsilon}{2}$, strictly worse than the expected value of choosing now.[23] Learning the outcome of the first coin flip can hurt Ann but it can't help her, so she is better off ignorant. Since Ann can foresee all of this, it is rational for her to reject free information. So the principle of maximizing expected utility sometimes recommends rejecting free information.

When I say that the information is 'free', I mean the same as Good: the information does not change the available actions or the utility function. The only impact of the information is to change Ann's credences. And it

---

[23]If Ann learns and decides afterwards, she chooses one of the risky options with probability ϵ and the safe option with probability $1-\epsilon$. The risky options have expected utility $-\frac{1}{2}$ and the safe option has expected utility zero. So the expected value of learning and deciding afterwards is $\epsilon \times -\frac{1}{2} + (1-\epsilon) \times 0 = -\frac{\epsilon}{2}$.

is not part of the example that Ann ever commits the gambler's fallacy. What makes it rational for Ann to reject the information is not that she actually deviates from conditionalization but that she assigns some positive probability to deviating from conditionalization.

I assume Ann is a 'sophisticated chooser': she predicts her own future choices and takes these predictions into account when making her present decisions (Hammond 1988, pp. 35–6).[24] Since she predicts that her future self might be irrational, she has an incentive to prevent her future self from making bad choices. So Ann faces a predicament similar to Odysseus in Greek mythology. She predicts that learning might compromise her future rationality, so she is better off ignorant.

You might complain that this example is a bit weird. Since Ann regards the two coin flips as independent, there is no way that learning the outcome of the first coin flip could help her make a better choice. At best, the information is neutral. In other words, if Ann is certain she will conditionalize, learning the outcome of the first coin flip is not relevant to her choice set. However, we can modify the example so that the information is relevant to her choice set but the principle of maximizing expected utility still recommends rejecting the information.

Again, a coin will be flipped twice and Ann must decide among several bets on the second coin flip. There is a safe bet, a slightly risky bet on heads, a slightly risky bet on tails, a very risky bet on heads and a very risky bet

---

[24]Buchak (2013, p. 176) describes the debate around sophisticated choice in decision theory. In moral philosophy, there is a similar debate between actualism and possibilism about what you should do when your future self will act wrongly (Smith 1976; Jackson and Pargetter 1986). Louise (2009) and White (2021) discuss the legitimate role of self-predictions in practical reasoning in more depth.

on tails in our choice set $\mathcal{S}$:

safe : {\$0 always},

heads : {\$1 if the second coin flip lands heads, $-$\$1 otherwise},

tails : {\$1 if the second coin flip lands tails, $-$\$1 otherwise},

v-risky-heads : {\$2 if the second coin flip lands heads, $-$\$10 otherwise},

v-risky-tails : {\$2 if the second coin flip lands tails, $-$\$10 otherwise}.

This time, Ann does not consider the coin to be fair but thinks that the coin has an unknown bias. The coin might be fair, it might be biased towards heads or it might be biased towards tails—she has no idea. Again, I offer Ann the following choice: She can either make her decision now or learn the outcome of the first coin flip and make her decision afterwards.

Since the coin has an unknown bias, observing the outcome of the first coin flip is informative for Ann. In particular, let us assume that, conditional on the first coin flip landing heads, Ann thinks that the second coin flip lands heads with probability $\frac{2}{3}$. The same goes for tails.[25]

If Ann makes her decision now, she is indifferent between safe, heads and tails. So the expected value of choosing now is:

$$\max_{f \in \mathcal{S}} \mathbb{E}_p(f) = \mathbb{E}_p(\texttt{safe}) = 0.$$

If Ann observes the outcome of the first coin flip, things are more interesting. Suppose Ann will conditionalize. Then if the coin lands heads, Ann will think that the coin is probably biased towards heads, so the slightly risky bet on heads will seem best to her. The very risky bet on heads will still seem too risky. The situation is analogous if the coin lands tails.

But Ann is modest and assigns some positive probability ϵ to the event that she overweights the evidence: when she learns that the first coin flip lands heads, she becomes .9 confident that the second coin flip will land heads and vice versa. So Ann takes the evidence into account, but thinks that she might be overconfident in how she does it. There are several reasons for why

---

[25]These probabilities can be derived from the 'rule of succession' (Zabell 1989).

Ann might do this. She might commit some version of the base rate fallacy, ignoring or underweighting prior probabilities. Or she might be susceptible to some form of the hot hand fallacy, believing that 'streaks' of successive heads are more likely than warranted by her evidence.

Suppose Ann observes the first coin flip landing heads. If she conditionalizes, she will take the slightly risky bet on heads. But if she is overconfident, she will choose the very risky bet on heads, which looks like a bad choice from her current point of view. A similar story applies if Ann observes the first coin flip landing tails. Figure 3 sums up Ann's situation.
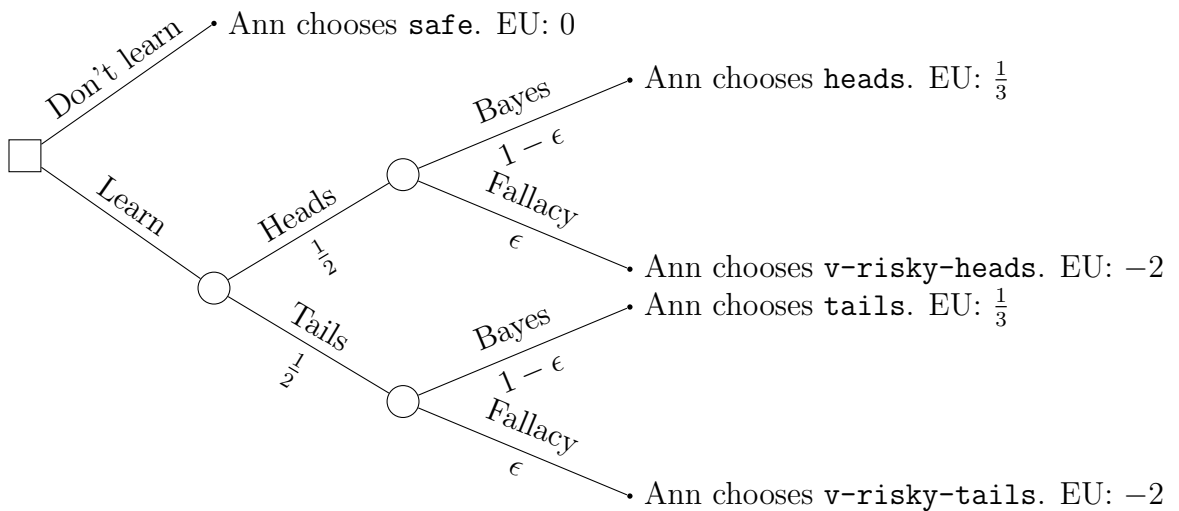

Figure 3: Ann's other decision problem.

The expected value of learning is $\frac{1}{3}(1 - \epsilon) - 2\epsilon$.[26] So if Ann thinks the probability of overconfidence is more than $\frac{1}{7}$, the expected value of learning and making her decision afterwards is worse than the expected value of deciding now.[27] So even if learning could be informative, expected utility maximizers can be required to reject free information. Again, it is not part

---

[26] If Ann learns and decides afterwards, she chooses one of the very risky options (v-risky-heads, v-risky-tails) with probability ϵ and one of the less risky options (heads, tails) with probability $1 - \epsilon$. The very risky options have expected utility -2 and the less risky options have expected utility $\frac{1}{3}$. So the expected value of learning and deciding afterwards is $\epsilon \times -2 + (1 - \epsilon) \times \frac{1}{3} = \frac{1}{3}(1 - \epsilon) - 2\epsilon$.

[27] Choosing now has expected value zero and $0 > \frac{1}{3}(1 - \epsilon) - 2\epsilon \iff \epsilon > \frac{1}{7}$.

of the example that Ann actually overweights her evidence but only that she assigns some probability to doing so.

These example show two things. First, the principle of maximizing expected utility does not imply Value of Learning. Expected utility maximizers are not always permitted to accept free and relevant information. Good's argument essentially depends on the assumption of Immodesty. If we assume Modesty, the principle of maximizing expected utility can require agents to reject free and relevant information.

Second, Value of Learning is false: Ann is rational but not permitted to accept free and relevant information. This might seem contentious. Whether or not rationality requires us to always conditionalize, the gambler's fallacy certainly looks irrational. So Ann thinks there is some probability that her future self will be irrational. However, the fact that Ann thinks her future self might be irrational does not entail that Ann is currently irrational. Rationality does not require you to be certain that your future self will be rational.

We can suppose that Ann has good evidence she might commit the gambler's fallacy. All her friends have committed it and she thinks they are relevantly similar to her. In this situation, it is implausible to think that Ann must be certain that her future self will conditionalize. Rather, if she is a good Bayesian, she should take her evidence into account and be modest. We can also suppose that Ann plans to conditionalize. Furthermore, we can suppose that in the actual word, Ann always manages to follow her plan. She just thinks that there is some chance she might fail. This does not seem like a failure of rationality. Therefore, we should let Value of Learning go even if we accept expected utility theory with precise credences and information which partitions logical space.

3.4   Information Aversion in the Real World

Moreover, we can use Modesty to make sense of real-world cases of information aversion. I will briefly illustrate this with medical testing, blind grading, checking one's stock portfolio and resisting manipulation.

People sometimes reject medical tests. There are several reasons: mistrust of doctors, fear of bad news and so on (Hertwig and Engel 2016, p. 365).[28] Modesty suggests another reason. People could fear that the test results might lead them (or their doctors) to draw inferences which they do not currently endorse, resulting in unnecessary treatment and further testing. For example, imagine you learn that a certain marker has increased in your blood test since last time but is still in the normal range. Learning this information might lead you to suspect a worrying trend where there are only random fluctuations. As a result, you might want another test soon which is unnecessary.

Blind grading is often considered good practice. Why is it bad to know the student's names? The standard explanation is that blind grading reduces bias. For example, I might give too much weight to the fact that George got an 'A' on the first paper and treat it as better evidence than it is that his current paper deserves a good grade.

It is sometimes suggested that you shouldn't check your stock portfolio daily. One reason is that it might stress you. However, another reason is that you might be tempted to change the allocation of your portfolio in a way that you currently view as a bad idea. This, in turn, can be explained by the risk of overweighting the significance of small fluctuations.[29]

It appears rational to refuse information designed to manipulate you. For example, one reason to avoid social media is that the information shown on your feed is designed to influence your behavior: to make you buy the products advertised there, to make you spend even more time on the platform and so on (Véliz 2020, pp. 69–76). Even if we sidestep issues of misinformation and assume the information you see on your feed is accurate, the fact that this information is designed to influence your behavior by companies which

---

[28]Information aversion with respect to medical tests is discussed by Osimani (2012), Jouini and Napp (2018) and many others.

[29]In a best-selling popular science book on computer science and decision theory, Christian and Griffiths (2016, p. 148) write: "If you want to be a good intuitive Bayesian—if you want to naturally make good predictions, without having to think about what kind of prediction rule is appropriate—you need to protect your priors. Counterintuitively, that might mean turning off the news". They do not consider how we can make sense of this idea without contradicting Good's theorem. Modesty offers an elegant way of doing so.

do not have your best interest at heart is a reason to stop looking. A similar case is when you refuse to talk to a manipulative person. Even if everything the manipulative person says is true, you might be better off not listening. This is because you might suspect that you will not update rationally on information designed to manipulate you.

Many other examples of information aversion in real life can be explained along similar lines.[30] There are, of course, competing explanations: perhaps people deviate from expected utility theory, have imprecise credences or assign negative utility to bad news. But in the examples above, it seems independently plausible that we assign some probability to overweighting evidence: giving too much weight to the result of medical tests, the past performance of our students, small fluctuations in our stock portfolio and the information on our feed. When we reflect on how much weight we should assign to this information, we might conclude: 'a little bit, but not very much'. But once we actually learn the information, we might assign more weight to it than we have previously considered rational. Imagine a positive result on a medical test slightly increases your probability of serious illness. Before doing the test, you might calmly assign conditional probabilities which reflect this slight increase. But when you learn that the test actually turned out positive, you might increase our probability of serious illness more than you previously considered warranted.

If we model overweighting evidence as deviation from conditionalization, we have seen that it can be a good idea to reject free information. So Modesty can explain these examples of information aversion in the real world in a way that seems to get at the heart of the matter. On the other hand, explaining these cases by saying, for example, that people are not expected utility maximizers seems to have less independent motivation. So while I have no conclusive argument that Modesty is the correct explanation for these cases, it seems like a particularly plausible candidate.[31]

---

[30]For example elite-group ignorance, which Kinney and Bright (2021) explain using risk-weighted expected utility theory. Yong (2023) critically discusses this explanation.

[31]Thanks to an anonymous referee for pushing me to clarify why Modesty is a plausible explanation of these cases of information aversion.

4   Value of Information Generalized

Surely, modest agents are not always required to reject free information. Even if you have some uncertainty about how you will update, this does not mean that you are always better off ignorant. But how should modest agents decide when to learn more information? And how general is the link between information aversion and Modesty? I answer these questions by generalizing the value of information to modest agents.

4.1   Good's Value of Information

Good's argument gives us a way to measure the value of information. To state this idea, it is useful to introduce an additional bit of notation. I write $\mathcal{P}(\cdot \mid \mathcal{E})$ for your credences updated by conditionalization on the evidence partition $\mathcal{E}$. This is a random variable which takes different probability functions as values in different state.[32] Then, we can define the value of information as follows (Blackwell 1951; Raiffa and Schlaifer 1961; Howard 1966):[33]

Definition 1. The value of information for $\mathcal{E}$ is:

$$Val_{Good}(\mathcal{E}) = \mathbb{E}_p\left(\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}(\cdot \mid \mathcal{E})}(f)\right) - \max_{f \in \mathcal{S}} \mathbb{E}_p(f).$$

This is the difference between the expected value of choosing after learning and the expected value of choosing now. It measures how much you expect the information to improve your decision. In this context, Value of Learning is captured by the fact that for any evidence partition $\mathcal{E}$, $Val_{Good}(\mathcal{E}) \geq 0$. In slogan form: 'the value of information is always non-negative'.

This concept is useful because it allows us to say how much you should value learning the answer to a question. It also allows us to compare the value

---

[32] More rigorously, we can define $\mathcal{P}(\cdot \mid \mathcal{E})$ as the random variable which maps each $\omega \in \Omega$ to $p(\cdot \mid E)$ for the unique $E \in \mathcal{E}$ such that $\omega \in E$.

[33] Le Cam (1996) sketches the history of this concept, which apparently goes back to an unpublished RAND memorandum entitled 'Reconnaissance in Game Theory' based on suggestions by von Neumann (Bohnenblust, Shapley, and Sherman 1949).

of learning the answers to different questions. This means we can formalize tradeoffs between acting now versus learning more and acting later even if learning is costly. Such tradeoffs are ubiquitous. In many real-life contexts, such as drug trials, we have to decide how much to sacrifice for learning more information.[34] So it is not surprising that the value of information is widely used in economics and artificial intelligence.[35] However, the standard way of defining the value of information presupposes Immodesty.

4.2 General Value of Information

Here is a proposal for how we can define the value of information in a more general way. I write $\mathcal{P}_{\mathcal{E}}$ for your credences updated on the evidence partition $\mathcal{E}$ without assuming you are certain you will update by conditionalization. This is a random variable whose values are different probability functions in different states.[36]

Suppose you will learn the true element of some evidence partition $\mathcal{E}$. Then you update your credences in some way—perhaps you conditionalize, perhaps you do something different—and choose the action which maximizes expected utility relative to your updated credences:

$$\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_{\mathcal{E}}}(f).^{37}$$

This expression will usually denote different actions in different states, because you might learn different events and update on those events in different

---

[34]Such decision problems can be formalized as 'multi-armed bandits' in which one must balance exploiting, acting according to one's current best estimate, and exploring new and potentially better options (Lattimore and Szepesvári 2020). The value of information can be used to define optimal solutions to such problems.

[35]Russell and Norvig (2018, pp. 628-33) discuss the value of information in AI research. Hadfield-Menell et al. (2017) discuss a model of how to ensure that AI agents always defer to humans which relies on the value of information being non-negative. In addition, the value of information is relevant to much other work, for example in the philosophy of language (Van Rooy 2003), to discussions about 'longtermism' in ethics (Askell and Neth Forthcoming) and the epistemology of disagreement (Dorst forthcoming).

[36]On each $E \in \mathcal{E}$, $\mathcal{P}_{\mathcal{E}}$ agrees with $\mathcal{P}_E$ as defined in section (2.1).

[37]The term $\arg\max_{x \in X} g(x)$ denotes the argument of the maximum: the $x \in X$ such that $g(x)$ is maximal.

ways. I assume that there is a unique best action in every state.[38]

We are interested in evaluating how good this action is from your current perspective, so we consider the expected utility of this action given your current credences:
$$\mathbb{E}_p\left(\arg\max_{f\in\mathcal{S}}\mathbb{E}_{\mathcal{P}_{\mathcal{E}}}(f)\right).$$

This is the expected utility of the action you think you will actually do after learning. We model a 'sophisticated chooser': our agent predicts her future choices and takes this information into account when making present decisions. I propose the following definition:

Definition 2. The general value of information for $\mathcal{E}$ is:

$$Val_{General}(\mathcal{E}) = \mathbb{E}_p\left(\arg\max_{f\in\mathcal{S}}\mathbb{E}_{\mathcal{P}_{\mathcal{E}}}(f)\right) - \max_{f\in\mathcal{S}}\mathbb{E}_p(f).$$

This measures the difference between the expected utility of your current best action and the expected utility of the action that you think you will choose after learning. In contrast to Good, I do not assume that you are certain you will conditionalize. I still assume you are certain you will maximize expected utility.

If we assume Immodesty, my proposal is identical to Good's:

Theorem 1. If $\mathcal{P}_E = p(\cdot \mid E)$ for all $E \in \mathcal{E}$, then $Val_{General}(\mathcal{E}) = Val_{Good}(\mathcal{E})$.

In contrast to Good's value of information, the general value of information can be negative. We have seen this in the examples above. But it is not always negative. This is shown by the second example above. When Ann considers the possibility of overconfidence sufficiently unlikely, she is better off observing the first coin flip. In (slightly clunky) slogan form: 'the value of information is sometimes negative, but not always. It depends'.

---

[38]So for every $\omega \in \Omega$, there is a unique $f^* \in \mathcal{S}$ maximizing $\mathbb{E}_{\mathcal{P}_{\mathcal{E}(\omega)}}(\cdot)$. Recall that $\mathcal{P}_{\mathcal{E}}$ is a function from states to probability functions, so $\mathcal{P}_{\mathcal{E}}(\omega)$ is a particular probability function—the credence you adopt after learning the true element of $\mathcal{E}$ in state ω. I do not consider cases where two actions are tied for the best action because in such cases, we would need to consider how to break the tie (introduce a selection function), which leads to additional complications. Buchak (2013, p. 190) provides relevant discussion and references on how indifference complicates sophisticated choice.

We can also say something about how general the link between Modesty and information aversion is. For this purpose, I make two additional assumptions. Utility Richness says that for every $x \in [0, 1]$, there is some outcome $o \in \mathscr{O}$ such that $u(o) = x$.[39] Evidential Independence says that conditional on the learned event, your updated credences are independent of what action is best.[40] Evidential Independence rules out cases where you deviate from conditionalizing because you become more certain of the truth. For example, you might observe that a fair coin lands heads and be able to foresee that it lands tails next. If we agree that your evidence is that the coin lands heads on the first flip, you do not update by conditionalizing on your evidence. However, clairvoyance can lead you to make better decisions than conditionalization. In contrast, I consider deviations from conditionalization which are not systematically correlated with which action is actually best. I have implicitly made this assumption earlier: Ann is equally likely to commit the gambler's fallacy whether the second coin flip lands heads or tails.

Assuming Evidential Independence, we can write $Val_{General}(\mathcal{E})$ as:

$$Val_{General}(\mathcal{E}) = \sum_{E \in \mathcal{E}} p(E) \sum_{i=1}^{n} p(\text{choose } f_i \mid E) \mathbb{E}_{p(\cdot|E)}(f_i) - \max_{f \in \mathcal{S}} \mathbb{E}_p(f),$$

where 'choose f_i' denotes the event that you choose action f_i after learning E, which means that f_i maximizes expected utility relative to your updated credences after learning E. Note that $p(\text{choose } f_i \mid E)$ is your current conditional probability that you will choose action f_i after learning E. You evaluate how good this action is by its conditional expected utility $\mathbb{E}_{p(\cdot|E)}(f_i)$ given your current credences. If you do not conditionalize, this conditional

---

[39]Our outcome space could contain lotteries which yield outcome b with probability p and outcome w with probability $(1-p)$. In this case, we only need two 'primitive' outcomes b and w with $u(b) > u(w)$ to obtain rich utilities.

[40]More precisely, for every $E \in \mathcal{E}$, your updated credences $\mathcal{P}_E$, which determine which action you will choose after learning, are independent of all $f \in \mathcal{S}$ conditional on E. This means, in particular, that for all $f, g \in S$, $\mathbb{E}_{p(\cdot|E\cap\text{choose } f)}(g) = \mathbb{E}_{p(\cdot|E)}(g)$, where 'choose f' is the event that you choose action f after learning. The intuition is that learning that you choose a particular action after learning E does not affect the expected utility of actions beyond learning E. It would be interesting to investigate cases where deviations from conditionalization are systematically correlated with which actions are best. In this paper, I focus on the simple case where Evidential Independence holds.

expected utility can come apart from the unconditional expected utility of the action according to your updated credences.[41]

We can show the following:

Theorem 2. Assuming Utility Richness and Evidential Independence, for every modest agent, there is some choice set where $Val_{General}(\mathcal{E}) < 0$.

Given our assumptions, any positive probability of not conditionalizing leads to information aversion. It does not matter why we are modest, as long as Evidential Independence holds. The examples above demonstrated how psychological uncertainty about whether you will update rationally leads to information aversion. The theorem shows that even if you are certain that your future self will be rational, normative uncertainty about whether conditionalization is rational leads to information aversion. So we have a very general argument against Value of Learning. This also means that we cannot rescue Good's argument by saying that while we might not be certain we will conditionalize, we are very confident we will conditionalize. (Perhaps we have a 'default entitlement' to believe in our future rationality.) Any non-zero probability of failing to conditionalize means trouble for Good.

We can also show:

Theorem 3. Assuming Evidential Independence, $\mathcal{E}$, $Val_{General}(\mathcal{E}) \leq Val_{Good}(\mathcal{E})$ for every evidence partition $\mathcal{E}$.

Doubts about how you will update cannot increase the value of information. Note that one might take this theorem as a reason to think that you should conditionalize, at least relative to the assumption of Evidential Independence. But we are often not sure whether we will be rational in the future and cannot change anything about that. If we are in such a situation, the theorem tells us that we should value learning less than if we were certain that we would conditionalize.

---

[41]Thanks to an anonymous referee for suggesting to make the formula for $Val_{General}(\mathcal{E})$ more explicit and see Lemma (1).

5   Conclusion

Good argues that the principle of maximizing expected utility entails Value of Learning: rational agents are always permitted to accept free information and required to accept information which is free and relevant. I have argued that the principle of maximizing expected utility does not entail Value of Learning and that Value of Learning is false. The key observation is that Good's argument only works if we are certain that we will update by conditionalization but we have good reason not to be.

What follows? First, we can give better advice to modest agents: sometimes, they are better off ignorant. Since we arguably are—and should be—modest, this advice applies to us. Sometimes, we are better off ignorant. Sometimes, we should avert our eyes and stuff our ears with wax to avoid learning the song of the Sirens. Second, proponents of expected utility theory should be careful when objecting to alternative decision-theoretic frameworks on the grounds that these frameworks sometimes permit or require agents to avoid free information. Properly understood, expected utility theory does the same. So this objection loses much of its dialectical force. Third, information aversion is a feature and not a bug. Plausible arguments from Bayesian epistemology push us towards Modesty. And once we accept Modesty, we can explain many instances of information aversion in the real world which would otherwise be puzzling. By going beyond Good, we end up with a better decision theory.

Appendix

Theorem 1. If $\mathcal{P}_E = p(\cdot \mid E)$ for all $E \in \mathcal{E}$, then $Val_{General}(\mathcal{E}) = Val_{Good}(\mathcal{E})$.

Proof. It suffices to show that if $\mathcal{P}_E = p(\cdot \mid E)$ for all $E \in \mathcal{E}$, then

$$\mathbb{E}_p \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_\mathcal{E}}(f) \right) = \mathbb{E}_p \left(\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}(\cdot|\mathcal{E})}(f) \right). \tag{1}$$

By the law of total expectation, since $\mathcal{E}$ is a partition with $p(E) > 0$ for all

$E \in \mathcal{E}$,[42]

$$\mathbb{E}_p \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_\mathcal{E}}(f) \right) = \sum_{E \in \mathcal{E}} p(E) \mathbb{E}_{p(\cdot | E)} \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}(f) \right).$$

We assume that $\mathcal{P}_E = p(\cdot \mid E)$ for all $E \in \mathcal{E}$, so

$$\sum_{E \in \mathcal{E}} p(E) \mathbb{E}_{p(\cdot | E)} \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}(f) \right) = \sum_{E \in \mathcal{E}} p(E) \mathbb{E}_{p(\cdot | E)} \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot | E)}(f) \right).$$

Now $\mathbb{E}_{p(\cdot | E)} \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot | E)}(f) \right) = \max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot | E)}(f)$, so

$$\sum_{E \in \mathcal{E}} p(E) \mathbb{E}_{p(\cdot | E)} \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot | E)}(f) \right) = \sum_{E \in \mathcal{E}} p(E) \max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot | E)}(f) = \mathbb{E}_p \left(\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}(\cdot | \mathcal{E})}(f) \right),$$

which shows that (1) holds. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 1. Assuming Evidential Independence, for every $f \in \mathcal{S}$,

$$\mathbb{E}_{p(\cdot | E)} \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}(f) \right) = \sum_{i=1}^{n} p(\text{choose } f_i \mid E) \mathbb{E}_{p(\cdot | E)}(f_i),$$

where 'choose f_i' is the event that you choose action f_i after learning E.

Proof. Suppose the range of $\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_\mathcal{E}}$ is $f_1, ..., f_n$. Intuitively, these are the actions you might choose after learning. Let us abbreviate the event $\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_\mathcal{E}} = f_i$ by 'choose f_i'. Intuitively, this is the event that you choose action f_i after learning E. (Recall that there is always a unique best action after learning.)

Evidential independence holds if for every $E \in \mathcal{E}$, $\mathcal{P}_E$ is independent of all $f \in \mathcal{S}$ conditional on E.[43] This means, in particular, that for all $f \in S$ and f_i with $1 \le i \le n$, $\mathbb{E}_{p(\cdot | E \cap \text{choose } f_i)}(f) = \mathbb{E}_{p(\cdot | E)}(f)$. The intuition is that

---

[42]In general, the law of total expectation says that for any random variables X and Y, $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X \mid Y))$ (Pitman 1993, p. 403). I use the special case where $\mathcal{E}$ is a partition with $p(E) > 0$ for all $E \in \mathcal{E}$ and X a random variable. Then $\mathbb{E}(X) = \sum_{E \in \mathcal{E}} p(E) \mathbb{E}(X \mid E)$. On every $E \in \mathcal{E}$, $\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_\mathcal{E}}$ agrees with $\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}$.

[43]Pitman (1993, p. 400) defines conditional independence for random variables.

relative to your prior, the event that you choose a particular action after learning E does not affect the expected utility of actions beyond learning E.

We want to show:

$$\mathbb{E}_{p(\cdot|E)}\left(\arg\max_{f\in\mathcal{S}}\mathbb{E}_{\mathcal{P}_E}(f)\right) = \sum_{i=1}^{n}p(\text{choose } f_i \mid E)\mathbb{E}_{p(\cdot|E)}(f_i). \qquad (2)$$

Since the events 'choose f_1', ..., 'choose f_n' form a partition, we can apply the law of total expectation:

$$\mathbb{E}_{p(\cdot|E)}\left(\arg\max_{f\in\mathcal{S}}\mathbb{E}_{\mathcal{P}_E}(f)\right) = \sum_{i=1}^{n}p(\text{choose } f_i \mid E)\mathbb{E}_{p(\cdot|E\cap\text{choose } f_i)}\left(\arg\max_{f\in\mathcal{S}}\mathbb{E}_{\mathcal{P}_E}(f)\right).$$

Now $\mathbb{E}_{p(\cdot|E\cap\text{choose } f_i)}\left(\arg\max_{f\in\mathcal{S}}\mathbb{E}_{\mathcal{P}_E}(f)\right) = \mathbb{E}_{p(\cdot|E\cap\text{choose } f_i)}(f_i)$ by the definition of 'choose f_i', so

$$\mathbb{E}_{p(\cdot|E)}\left(\arg\max_{f\in\mathcal{S}}\mathbb{E}_{\mathcal{P}_E}(f)\right) = \sum_{i=1}^{n}p(\text{choose } f_i \mid E)\mathbb{E}_{p(\cdot|E\cap\text{choose } f_i)}(f_i).$$

By Evidential Independence, $\mathbb{E}_{p(\cdot|E\cap\text{choose } f_i)}(f_i) = \mathbb{E}_{p(\cdot|E)}(f_i)$, so (2) holds. $\qquad\square$

Theorem 2. Assuming Utility Richness and Evidential Independence, for every modest agent, there is some choice set where $Val_{General}(\mathcal{E}) < 0$.

Proof. An agent is modest iff she assigns some positive probability to not conditionalizing. So for some evidence partition $\mathcal{E}$, there is some $E \in \mathcal{E}$ such that with positive probability, $\mathcal{P}_E \neq p(\cdot \mid E)$. This means that for some $\omega \in E$ with $p(\omega) > 0$, $\mathcal{P}_E(\omega)(A) \neq p(A \mid E)$ for some event A. I write p_E for $\mathcal{P}_E(\omega)$ and assume $p_E(A) > p(A \mid E)$. (In the other case, the proof is analogous.)

We want to show that there is a choice set where $Val_{General}(\mathcal{E}) < 0$. Consider the following choice set $\mathcal{S}$ (with payoffs in utils):

$$\texttt{safe} : \{0 \text{ always}\},$$
$$\texttt{risky} : \{a \text{ if } A \cap E, -b \text{ if } A^C \cap E, 0 \text{ otherwise}\}.$$

We want to find values for $a > 0$ and $b > 0$ such that, conditional on E, the expected utility of the risky bet is worse:

$$\mathbb{E}_{p(\cdot|E)}(\texttt{risky}) < 0. \tag{3}$$

But if our agent deviates from conditionalization, she prefers the risky bet:

$$\mathbb{E}_{p_E}(\texttt{risky}) > 0. \tag{4}$$

If we find these values, we can show that $Val_{General}(\mathcal{E}) < 0$. Recall that

$$Val_{General}(\mathcal{E}) = \mathbb{E}_p\left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_\mathcal{E}}(f)\right) - \max_{f \in \mathcal{S}} \mathbb{E}_p(f).$$

Now $\max_{f \in \mathcal{S}} \mathbb{E}_p(f) = 0$. This is because $\mathbb{E}_p(\texttt{safe}) = 0$ but $\mathbb{E}_p(\texttt{risky}) < 0$. We need to show

$$\mathbb{E}_p\left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_\mathcal{E}}(f)\right) < 0. \tag{5}$$

We re-write this term using the law of total expectation:

$$p(E)\mathbb{E}_{p(\cdot|E)}\left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}(f)\right) + p(E^C)\mathbb{E}_{p(\cdot|E^C)}\left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_{E^C}}(f)\right).$$

Now the right-hand term is zero, since both $\texttt{safe}$ and $\texttt{risky}$ yield zero when E is false. Thus, we focus on the left-hand term, which we can re-write as follows, using Evidential Independence and Lemma (1):

$$p(\text{choose } \texttt{risky} \mid E)\mathbb{E}_{p(\cdot|E)}(\texttt{risky}) + p(\text{choose } \texttt{safe} \mid E)\mathbb{E}_{p(\cdot|E)}(\texttt{safe}).$$

The right-hand term is again zero, so we focus on the left-hand term. We have $p(\text{choose } \texttt{risky} \mid E) > 0$, since we have assumed that there is a positive probability our agent deviates from conditionalization and so chooses the risky action. By assumption, $\mathbb{E}_{p(\cdot|E)}(\texttt{risky}) < 0$, which shows (5).

We still need to show that we can find values a and b which do the trick.

By (3), a and b need to obey the following constraint:

$$ap(A \cap E \mid E) - bp(A^C \cap E \mid E) < 0,$$

which simplifies to

$$ap(A \mid E) - b(1 - p(A \mid E)) < 0. \tag{6}$$

By (4), a and b need to obey the following constraint:

$$ap_E(A \cap E) - bp_E(A^C \cap E) > 0,$$

which, using our assumption that $p_E(E) = 1$, simplifies to

$$ap_E(A) - b(1 - p_E(A)) > 0. \tag{7}$$

Let us write q for $p_E(A)$ and r for $p(A \mid E)$. So our question is whether the following system of equations has a solution for any q and r such that $q > r$:

$$aq - b(1 - q) > 0 > ar - b(1 - r). \tag{8}$$

The answer is 'yes': real numbers $a > 0$ and $b > 0$ such that $0 \le r < \frac{b}{a+b} < q \le 1$. We can find outcomes with these utilities by Utility Richness. $\square$

Theorem 3. Assuming Evidential Independence, $\mathcal{E}$, $Val_{General}(\mathcal{E}) \le Val_{Good}(\mathcal{E})$ for every evidence partition $\mathcal{E}$.

Proof. It suffices to show

$$\sum_{E \in \mathcal{E}} p(E)\mathbb{E}_{p(\cdot \mid E)}\left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}(f)\right) \le \mathbb{E}_p\left(\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}(\cdot \mid \mathcal{E})}(f)\right). \tag{9}$$

Consider any $E \in \mathcal{E}$. By Evidential Independence and Lemma (1),

$$\mathbb{E}_{p(\cdot \mid E)}\left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}(f)\right) = \sum_{i=1}^{n} p(\text{choose } f_i \mid E)\mathbb{E}_{p(\cdot \mid E)}(f_i). \tag{10}$$

The right-hand side is a weighted sum of expected values relative to $p(\cdot \mid E)$ and $\mathbb{E}_{p(\cdot|E)}(f) \leq \max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot|E)}(f)$ for all $f \in S$.[44] Therefore,

$$\sum_{i=1}^{n} p(\text{choose } f_i \mid E) \mathbb{E}_{p(\cdot|E)}(f_i) \leq \max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot|E)}(f),$$

and so by (10),

$$\mathbb{E}_{p(\cdot|E)} \left(\arg\max_{f \in \mathcal{S}} \mathbb{E}_{\mathcal{P}_E}(f) \right) \leq \max_{f \in \mathcal{S}} \mathbb{E}_{p(\cdot|E)}(f).$$

Taking expectations on both sides, (9) follows. $\square$

Acknowledgements

---

[44]This is a version of the principle that, for expected utility maximizers, randomization can never be strictly preferable (Icard 2021, pp. 119–120). In general, this follows from Jensen's inequality.

References

Adams, Ernest W. and Roger D. Rosenkrantz (1980). "Applying the Jeffrey Decision Model to Rational Betting and Information Acquisition". In: *Theory and Decision* 12.1, pp. 1–20. DOI: `10.1007/BF00154655`.

Ahmed, Arif (2016). "Lara Buchak's Risk and Rationality". In: *BJPS Review of Books*.

Al-Najjar, Nabil I. and Jonathan Weinstein (2009). "The Ambiguity Aversion Literature: A Critical Assessment". In: *Economics & Philosophy* 25.3, pp. 249–284. DOI: `10.1017/S026626710999023X`.

Askell, Amanda and Sven Neth (Forthcoming). "Longtermist Myopia". In: *Essays on Longtermism*. Ed. by Hilary Greaves Jacob Barrett and David Thorstad. Oxford: Oxford University Press.

Blackwell, David (1951). "Comparison of Experiments". In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Vol. 2. Berkeley and Los Angeles: University of California Press, pp. 93–103.

Bohnenblust, Henri Frédéric, Lloyd S. Shapley, and Seymour Sherman (1949). *Reconnaissance in Game Theory*. Rand Memorandum RM. 208.

Bradley, Darren (2020). "Bayesianism and Self-Doubt". In: *Synthese* 199.1-2, pp. 2225–2243. DOI: `10.1007/s11229-020-02879-7`.

Bradley, Seamus and Katie Steele (2016). "Can Free Evidence Be Bad? Value of Information for the Imprecise Probabilist". In: *Philosophy of Science* 83.1. DOI: `10.1086/684184`.

Bratman, Michael E. (1992). "Practical Reasoning and Acceptance in a Context". In: *Mind* 101.401, pp. 1–16. DOI: `10.1093/mind/101.401.1`.

Briggs, Ray (2009). "Distorted Reflection". In: *Philosophical Review* 118.1, pp. 59–85. DOI: `10.1215/00318108-2008-029`.

— (2015). "Costs of Abandoning the Sure-Thing Principle". In: *Canadian Journal of Philosophy* 45.5, pp. 827–840. DOI: `10.1080/00455091.2015.1122387`.

Brown, Peter M. (1976). "Conditionalization and Expected Utility". In: *Philosophy of Science* 43.3, pp. 415–419. DOI: `10.1086/288696`.

Buchak, Lara (2010). "Instrumental Rationality, Epistemic Rationality, and Evidence-Gathering". In: *Philosophical Perspectives* 24.1, pp. 85–120. DOI: 10.1111/j.1520-8583.2010.00186.x.

— (2013). *Risk and Rationality*. Oxford: Oxford University Press.

Campbell-Moore, Catrin and Bernhard Salow (2020). "Avoiding Risk and Avoiding Evidence". In: *Australasian Journal of Philosophy* 98.3, pp. 495–515. DOI: 10.1080/00048402.2019.1697305.

Carr, Jennifer Rose (2019). "A Modesty Proposal". In: *Synthese* 198.4, pp. 3581–3601. DOI: 10.1007/s11229-019-02301-x.

Christensen, David (2007). "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals". In: *Oxford Studies in Epistemology: Volume 2*. Ed. by Tamar Szabo Gendler and John Hawthorne. Oxford: Oxford University Press, pp. 3–31.

Christian, Brian and Tom Griffiths (2016). *Algorithms To Live By: The Computer Science of Human Decisions*. New York: Henry Holt.

Cohen, Michael (2020). "Opaque Updates". In: *Journal of Philosophical Logic* 50.3, pp. 447–470. DOI: 10.1007/s10992-020-09571-8.

Das, Nilanjan (2023). "The Value of Biased Information". In: *British Journal for the Philosophy of Science* 74.1, pp. 25–55. DOI: 10.1093/bjps/axaa003.

Diaconis, Persi and Sandy L. Zabell (1982). "Updating Subjective Probability". In: *Journal of the American Statistical Association* 77.380, pp. 822–830. DOI: 10.1080/01621459.1982.10477893.

Dorst, Kevin (2020). "Evidence: A Guide for the Uncertain". In: *Philosophy and Phenomenological Research* 100.3, pp. 586–632. DOI: 10.1111/phpr.12561.

— (forthcoming). "Rational Polarization". In: *The Philosophical Review*.

Douven, Igor (2013). "Inference to the Best Explanation, Dutch Books, and Inaccuracy Minimisation". In: *Philosophical Quarterly* 63.252, pp. 428–444. DOI: 10.1111/1467-9213.12032.

Golman, Russell, David Hagmann, and George Loewenstein (2017). "Information Avoidance". In: *Journal of Economic Literature* 55.1, pp. 96–135. DOI: 10.1257/jel.20151245.

Good, Irving John (1967). "On the Principle of Total Evidence". In: *British Journal for the Philosophy of Science* 17.4, pp. 319–321. DOI: `10.1093/bjps/17.4.319`.

— (1974). "A Little Learning Can Be Dangerous". In: *British Journal for the Philosophy of Science* 25.4, pp. 340–342. DOI: `10.1093/bjps/25.4.340`.

Greaves, Hilary and David Wallace (2005). "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility". In: *Mind* 115.459, pp. 607–632. DOI: `10.1093/mind/fzl607`.

Gyenis, Zalán and Miklós Rédei (2017). "General Properties of Bayesian Learning as Statistical Inference Determined by Conditional Expectations". In: *Review of Symbolic Logic* 10.4, pp. 719–755. DOI: `10.1017/s1755020316000502`.

Hacking, Ian (1967). "Slightly More Realistic Personal Probability". In: *Philosophy of Science* 34.4, pp. 311–325. DOI: `10.1086/288169`.

Hadfield-Menell, Dylan, Anca Dragan, Pieter Abbeel, and Stuart Russell (2017). "The Off-Switch Game". In: *IJCAI-17: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. Ed. by Carles Sierra, pp. 220–227. DOI: `10.24963/ijcai.2017`.

Hammond, Peter J. (1988). "Consequentialist Foundations for Expected Utility". In: *Theory and Decision* 25.1, pp. 25–78. DOI: `10.1007/bf00129168`.

Hertwig, Ralph and Christoph Engel (2016). "Homo Ignorans: Deliberately Choosing Not To Know". In: *Perspectives on Psychological Science* 11.3, pp. 359–372. DOI: `10.1177/1745691616635594`.

Hosiasson, Janina (1931). "Why Do We Prefer Probabilities Relative to Many Data?" In: *Mind* 40.157, pp. 23–36. DOI: `10.1093/mind/xl.157.23`.

Howard, Ronald (1966). "Information Value Theory". In: *IEEE Transactions on Systems Science and Cybernetics* 2.1, pp. 22–26. DOI: `10.1109/TSSC.1966.300074`.

Huttegger, Simon M. (2014). "Learning Experiences and the Value of Knowledge". In: *Philosophical Studies* 171.2, pp. 279–288. DOI: `10.1007/s11098-013-0267-7`.

Icard, Thomas (2021). "Why Be Random?" In: *Mind* 130.517, pp. 111–139. DOI: `10.1093/mind/fzz065`.

Jackson, Frank and Robert Pargetter (1986). "Oughts, Options, and Actualism". In: *Philosophical Review* 95.2, pp. 233–255. DOI: 10.2307/2185591.

James, Henry ([1882] 2011). *The Portrait of a Lady*. London: Penguin.

Jouini, Elyès and Clotilde Napp (2018). "The Impact of Health-Related Emotions on Belief Formation and Behavior". In: *Theory and Decision* 84.3, pp. 405–427. DOI: 10.1007/s11238-017-9610-3.

Joyce, James (1998). "A Nonpragmatic Vindication of Probabilism". In: *Philosophy of Science* 65.4, pp. 575–603. DOI: 10.1086/392661.

Kadane, Joseph B., Mark Schervish, and Teddy Seidenfeld (1996). "Reasoning to a Foregone Conclusion". In: *Journal of the American Statistical Association* 91.435, pp. 1228–1235. DOI: 10.1080/01621459.1996.10476992.

— (2008). "Is Ignorance Bliss?" In: *Journal of Philosophy* 105.1, pp. 5–36. DOI: 10.5840/jphil200810518.

Kahneman, Daniel and Amos Tversky (1973). "On the Psychology of Prediction". In: *Psychological Review* 80.4, pp. 237–251. DOI: 10.1037/h0034747.

Kinney, David and Liam Kofi Bright (2021). "Risk Aversion and Elite-Group Ignorance". In: *Philosophy and Phenomenological Research* 106.1, pp. 35–57. DOI: 10.1111/phpr.12837.

Laffont, Jean-Jacques (1989). *The Economics of Uncertainty and Information*. Boston: MIT Press.

Lattimore, Tor and Csaba Szepesvári (2020). *Bandit Algorithms*. Cambridge: Cambridge University Press.

Le Cam, Lucien (1996). "Comparison of Experiments: A Short Review". In: *Statistics, Probability and Game Theory: Papers in Honor of David Blackwell*. Vol. 30. Institute of Mathematical Statistics Lecture Notes - Monograph Series. Institute of Mathematical Statistics, pp. 127–138. DOI: 10.1214/lnms/1215453569.

Lederman, Harvey (2015). "People with Common Priors Can Agree to Disagree". In: *Review of Symbolic Logic* 8.1, pp. 11–45. DOI: 10.1017/s1755020314000380.

Lewis, David (1980). "A Subjectivist's Guide to Objective Chance". In: *Studies in Inductive Logic and Probability, Volume II*. Ed. by Richard C. Jeffrey. Berkeley and Los Angeles: University of California Press, pp. 263–293.

— (1999). "Why Conditionalize". In: *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press, pp. 403–407.

Lindley, Dennis V. (1965). *Introduction to Probability and Statistics from a Bayesian Point of View. Part 2: Inference*. Cambridge: Cambridge University Press.

Louise, Jennie (2009). "I Won't Do It! Self-Prediction, Moral Obligation and Moral Deliberation". In: *Philosophical Studies* 146.3, pp. 327–348. DOI: `10.1007/s11098-008-9258-5`.

Maher, Patrick (1990). "Symptomatic Acts and the Value of Evidence in Causal Decision Theory". In: *Philosophy of Science* 57.3, pp. 479–498. DOI: `10.1086/289569`.

Osimani, Barbara (2012). "Risk Information Processing and Rational Ignoring in the Health Context". In: *The Journal of Socio-Economics* 41.2, pp. 169–179. DOI: `10.1016/j.socec.2011.10.009`.

Paul, Laurie Ann (2014). *Transformative Experience*. Oxford: Oxford University Press.

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

— (2019). *Choosing for Changing Selves*. Oxford: Oxford University Press.

— (2020). "What is Conditionalization, and Why Should We Do It?" In: *Philosophical Studies* 177.11, pp. 3427–3463. DOI: `10.1007/s11098-019-01377-y`.

Pitman, Jim (1993). *Probability*. New York: Springer.

Raiffa, Howard and Robert Schlaifer (1961). *Applied Statistical Decision Theory*. Boston: Graduate School of Business Administration, Harvard University.

Ramsey, Frank (1931). "Truth and Probability". In: *The Foundations of Mathematics and other Logical Essays*. Ed. by R.B. Braithwaite. 1999

electronic edition. New York: Harcourt, pp. 156–198. DOI: 10.4324/9781315887814.

— (1990). "Weight or the Value of Knowledge". In: *British Journal for the Philosophy of Science* 41.1, pp. 1–4. DOI: 10.1093/bjps/41.1.1.

Russell, Stuart and Peter Norvig (2018). *Artificial Intelligence: A Modern Approach*. Third Edition. Hoboken, NJ: Prentice Hall.

Salow, Bernhard and Arif Ahmed (2019). "Don't Look Now". In: *British Journal for the Philosophy of Science* 70.2, pp. 327–350. DOI: 10.1093/bjps/axx047.

Savage, Leonard J. (1972). *The Foundations of Statistics*. Second Revised Edition. Hoboken, NJ: Wiley Publications in Statistics.

Skyrms, Brian (1990). "The Value of Knowledge". In: *Scientific Theories*. Ed. by C.W. Savage. Vol. 14. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, pp. 245–266.

Smith, Holly (1976). "Dated Rightness and Moral Imperfection". In: *Philosophical Review* 85.4, pp. 449–487. DOI: 10.2307/2184275.

Van Rooy, Robert (2003). "Quality and Quantity of Information Exchange". In: *Journal of Logic, Language and Information* 12, pp. 423–451. DOI: 10.1023/A:1025054901745.

Véliz, Carissa (2020). *Privacy Is Power*. London: Penguin (Bantam Press).

Wakker, Peter (1988). "Nonexpected Utility As Aversion of Information". In: *Journal of Behavioral Decision Making* 1.3, pp. 169–175. DOI: 10.1002/bdm.3960010305.

Wheeler, Gregory (2021). "Less is More for Bayesians, Too." In: *Routledge Handbook on Bounded Rationality*. Ed. by Riccardo Viale. London and New York: Routledge, pp. 471–483.

White, Stephen J. (2021). "Self-Prediction in Practical Reasoning: Its Role and Limits". In: *Noûs* 55.4, pp. 825–841. DOI: 10.1111/nous.12333.

Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Yong, Xin Hui (2023). "Risk, Rationality and (Information) Resistance: De-Rationalizing Elite-Group Ignorance". In: *Erkenntnis*. DOI: 10.1007/s10670-022-00656-1.

Zabell, Sandy L. (1989). "The Rule of Succession". In: *Erkenntnis* 31.2-3, pp. 283–321. DOI: 10.1007/BF01236567.