⋆

# A Bayesian account of establishing

⋆

Jon Williamson

April 20, 2020
*British Journal for the Philosophy of Science*

**Abstract**

When a proposition is established, it can be taken as evidence for other propositions. Can the Bayesian theory of rational belief and action provide an account of establishing? I argue that it can, but only if the Bayesian is willing to endorse objective constraints on both probabilities and utilities, and willing to deny that it is rationally permissible to defer wholesale to expert opinion. I develop a new account of deference that accommodates this latter requirement.

## §1 Introduction

This paper asks whether Bayesianism can provide an account of establishing. When a proposition is established, it can be taken as evidence for other propositions. Since Bayesianism is intended to explicate the relation between evidence on the one hand and rational beliefs and decisions on the other, it is essential that Bayesianism is able to accommodate the generation of new evidence.

This question is distinct from the question of whether Bayesianism can provide an account of acceptance. As Kaplan (1981) pointed out, acceptance is not a state of certainty: one can accept a proposition as a basis for action without fully believing the proposition thereafter. For example, one can accept Newton's laws of motion when performing some simple engineering task, yet it is reasonable to believe these laws to a very low degree—even to degree zero—given that they have been falsified and superseded by more complex laws. When a proposition is established, on the other hand, one ought to fully believe it: it is a core tenet of Bayesianism that one ought to fully believe one's evidence. Therefore, the literature on rational acceptance is largely concerned with a different phenomenon to that considered here. A similar point can be made about the literature on the relationship between qualitative belief and degree of belief: we would be reasonable to give many of our qualitative beliefs credence less than 1, if pressed. My (qualitative) belief that it will rain in Canterbury in the next week is well justified, yet it would be reasonable to give this proposition credence less than 1 and it would not be appropriate for me to treat it as evidence. Hence, believing is also distinct from establishing.

The paper is structured as follows. §2 explains the problem, as well as the scope and presuppositions of this paper. §3 shows that the Bayesian account of establishing is essentially decision-theoretic. §4 argues that, to be viable, this account must admit objective constraints on rational degrees of belief. §5 argues that it must also admit substantive constraints on utilities. §6 argues that a viable account of establishing must also avoid wholesale deference to expert beliefs, and develops a new account of discounted deference. §7 takes stock and considers some potential objections. Appendices 1 and 2 show that the analyses of §3 and §5 continue to hold in a more complex decision-making setting.

## §2
## Bayesianism, evidence and establishing

This paper will remain neutral about the nature of evidence. Some suggest that your evidence is what you know (Williamson, 2000), others that your evidence is constituted by your true beliefs (Mitova, 2017), or your credences that are set by observation (Jeffrey, 2004), or your information (Rowbottom, 2014), or what you rationally grant (Williamson, 2015). Each of these theories will have something to say about what it is to establish a proposition; however, the claims made in this paper are independent of which of these theories is true, if any.

Bayesianism is concerned with the questions of when the strengths of an agent's beliefs, and the decisions informed by these degrees of belief, are rational. The answer it provides to the latter question is known as Bayesian decision theory and is sketched in §3, while the answer it provides to the former question is known as Bayesian probability theory or Bayesian epistemology and is described in §4.

There are two perspectives one might have on these questions. The usual perspective is that of the deliberating agent herself, interested in determining how strongly to believe certain propositions and which acts to choose in certain decision problems. Under this actor's perspective, Bayesianism is a practical theory; 'to us, probability is the very guide of life' (Butler, 1736, p.xxv). An alternative perspective is purely theoretical: concerned with what it takes for certain degrees of belief and decisions to be rational, but not engaged with determining rational degrees of belief or making decisions.

On either perspective, degrees of belief are deemed rational if they are appropriate given the evidence. The various Bayesian norms, described in §4, explicate what it is for degrees of belief to be appropriate given the evidence. The key difference between the two perspectives is that on the former, practical perspective, the norms had better be followable and your evidence accessible, for otherwise Bayesianism could offer little guidance. On the latter, theoretical perspective, there may be no such requirement: some purely theoretical development of Bayesianism might deem your degrees of belief to be irrational even if you are in no position to determine your evidence, nor to follow Bayesian norms.

The theoretical perspective might presuppose an externalist account of evidence, for example. Then we might distinguish your evidence in this externalist sense from what you take to be your evidence (which we shall sometimes call 'wyttby evidence' for short). From the practical perspective, your degrees of belief are rational iff they conform to the Bayesian norms given wyttby evidence, as long as you are rational to take as your evidence what you take to be your evidence and you are in a position to follow the Bayesian norms.

This paper will adopt the usual practical perspective—Bayesianism as a guide to life. For this reason, we will focus on wyttby evidence, rather than some more theoretical notion of evidence.

For the purposes of this paper, we can be flexible about the kinds of propositions that can constitute rational wyttby evidence. In addition to past observations, evidence might include reasonable presuppositions, universal laws and even statements about the future. For example, you might take a proposition about the timing of tomorrow's tides as evidence that bears on your decision as to when to forage for oysters. Butler has something to say about when to take such a proposition as a 'moral certainty':

> Probable evidence is essentially distinguished from demonstrative by this, that it admits of degrees; and of all variety of them, from the highest moral certainty, to the very lowest presumption. We cannot indeed say a thing is probably true upon one very slight presumption for it; because, as there may be probabilities on both sides of a question, there may be some against it; and though there be not, yet a slight presumption does not beget that degree of conviction, which is implied in saying a thing is probably true. But that the slightest possible presumption is of the nature of a probability, appears from hence; that such low presumption often repeated, will amount even to moral certainty. Thus a man's having observed the ebb and flow of the tide to-day, affords some sort of presumption, though the lowest imaginable, that it may happen again to-morrow: but the observation of this event for so many days, and months, and ages together, as it has been observed by mankind, gives us a full assurance that it will. (Butler, 1736, p.xxiv)

Butler's discussion highlights two key features of establishing. First, in order to establish a proposition, strength of belief in that proposition must meet a substantial threshold. After all, the proposition needs to become evident. Second, there must be enough evidence to make this degree of belief sufficiently stable—to establish is to render stable or firm. We will refer to these two key features as *threshold* and *stability* respectively.[1]

One can also distinguish between establishing as an act and establishing as an evidential relation. Where establishing is an act, it is a subject $S$ that establishes proposition $A$. Where establishing is an evidential relation, it is a set $\boldsymbol{E}$ of propositions that establishes proposition $A$.

Here is an example of the act of establishing: CERN established in 2012 that they had observed a new particle in the mass region around 125 GeV. It is not always obvious whether a statement refers to the act of establishing or to the evidential relation. For example, 'It is not yet established that this particle was the Higgs boson', might be read either as a statement about the act of establishing (no one has yet established that this was the Higgs boson, nor is anyone yet in a position to do so) or as a statement about the evidential relation (current evidence does not establish that this was the Higgs boson). Plausibly, this ambiguity exists because the act and the evidential relation are closely connected. One might think of the act of establishing $A$ as the act of taking $A$ to be established by wyttby evidence,

[1] Skyrms (1977) offers a formal explication of stability, which he calls 'resiliency'. We do not need to commit to any particular formal notion of stability here: all that is required is the observation that establishing $A$ may be reasonable when one's degree of belief in $A$ reaches a high threshold and is likely to remain above that threshold, but unreasonable should one's degree of belief be very volatile over time and unlikely to remain above the high threshold. See Parkkinen et al. (2018, §3.2) for discussion of threshold and stability in the context of establishing causal claims in medicine, and Leitgeb (2017) for a discussion of stability in relation to the connection between degree of belief and qualitative belief.

in which case you are rational to establish $A$ iff wyttby evidence establishes $A$. On the other hand, one might argue that $E$ establishes $A$ iff taking $E$ as evidence puts you in a position to establish $A$.

Given the perspective on Bayesianism taken in this paper—using wyttby evidence as a guide to life—the act of establishing will be of primary interest.

Clearly, a contingent proposition cannot be established by evidence that has no bearing whatsoever on its truth. This platitude extends to the act of establishing, in virtue of the connection between the act and the evidential relation: it is not rationally permissible to establish a contingent proposition in the absence of evidence that bears on its truth. We will refer to this feature of establishing as ***non-vacuity***. Non-vacuity does not imply that one cannot take any proposition as evidence for purely pragmatic reasons—merely that such a proposition cannot be said to be established by prior evidence. To assume that all evidence must be established by prior evidence would lead to a kind justificatory regress, and no such assumption will be made here. We will return to non-vacuity in §4.

The final observation to make is that the act of establishing is fallible. For example, it is quite possible that although CERN established in 2012 that they had observed a new particle, as it turns out, they were mistaken. There are various ways in which such an act of establishing might fail. Perhaps, wyttby evidence $E$ at that time did not, after all, establish the existence of the new particle. For example, perhaps CERN were irrational to take existence of the particle to be established by what they took to be their evidence because their statistical analysis was flawed. Or perhaps they were irrational to take $E$ as evidence in the first place. Externalism about evidence opens up a further possibility: that CERN were rational to take $E$ as evidence, but as a matter of fact $E$ was erroneous and did not correspond to their evidence in the externalist sense, which did not establish existence. In which case, CERN were rational to establish existence from the practical Bayesian perspective, but not from the theoretical perspective. A proponent of a theoretical perspective might even deem the evidential relation of establishing to be factive. Then a fourth possibility arises: $E$ qualifies as evidence in the externalist sense and strongly confirms the existence of the particle, but as a matter of fact the particle does not exist.

Given the practical perspective adopted here, we shall focus on the first of these four kinds of error. Assuming that it is rational to take wyttby evidence $E$ as evidence, we shall ask when it is rational to establish a proposition $A$ on the basis of $E$. In the next section we see that Bayesianism provides a precise answer to this question. In subsequent sections I argue that this answer imposes strong constraints on Bayesianism itself.

## §3
## A decision-theoretic criterion

Since the act of establishing is an act, it falls under the remit of Bayesian decision theory. Bayesian decision theory says that you should choose, out of a range of acts, one that maximises expected utility. The probabilities used to calculate the expected utility of each act are your degrees of belief in relevant states of the world.

The first step is to determine the utility of each act in each state of the world. In the simplest case of establishing, there are two acts—establish $A$, don't establish

$A$—and the utility of each act depends on whether $A$ is true or false:[2]

| | $A$ | $\bar{A}$ |
|---|---|---|
| establish $A$ | $s_A$ | $e_{\bar{A}}$ |
| don't establish $A$ | $e_A$ | $s_{\bar{A}}$ |

Here $s_A$ is the utility of successfully establishing $A$, i.e., of establishing $A$ when $A$ is true; $e_A$ is the utility of erroneously failing to establish $A$, i.e., of not establishing $A$ when $A$ is true; $s_{\bar{A}}$ is the utility of successfully not establishing $A$, i.e., of not establishing $A$ when $A$ is false; and $e_{\bar{A}}$ is the utility of erroneously establishing $A$, i.e., of establishing $A$ when $A$ is false. It is important to keep in mind that the act of not establishing $A$ is different to the act of establishing $\bar{A}$. Deciding, on the basis of wyttby evidence, not to establish the existence of a new particle in the mass region around 125 GeV is not the same as deciding to establish that there is no new particle in that mass region.

Bayesian decision theory says that you ought to establish $A$ iff the expected utility of establishing $A$ is greater than that of not establishing $A$:

$$P(A)s_A + P(\bar{A})e_{\bar{A}} > P(A)e_A + P(\bar{A})s_{\bar{A}},$$

where $P$ is your personal probability function, which captures your rational degrees of belief. This condition is met just when

$$P(A) > \frac{s_{\bar{A}} - e_{\bar{A}}}{(s_A - e_A) + (s_{\bar{A}} - e_{\bar{A}})},$$

since $P(\bar{A}) = 1 - P(A)$. The right-hand side of this inequality defines a threshold above which $A$ ought to be established. We shall denote this threshold by $\tau_A$. $s_A - e_A$ is the advantage in establishing $A$ when $A$ is true; call this the 'true advantage'. $s_{\bar{A}} - e_{\bar{A}}$ is the advantage in not establishing $A$ when $A$ is false; call this the 'false advantage'. Then,

$$\tau_A = \frac{\text{false advantage}}{\text{true advantage + false advantage}}$$

While Bayesian decision theory deems establishing $A$ to be rationally required when $P(A) > \tau_A$, it deems establishing $A$ to be rationally permissible when $P(A) \geq \tau_A$. If $P(A) = \tau_A$, not establishing $A$ is also rationally permissible, as the two options have the same expected utility. If $P(A) < \tau_A$, establishing $A$ is rationally impermissible.[3] The criterion for establishing $A$ can be expressed in terms of odds: you are rationally required to establish $A$ iff your odds for $A$ exceed the ratio of the false advantage to the true advantage.

In sum, the decision-theoretic aspects of Bayesianism impose a particular criterion for establishing $A$. Thus Bayesianism offers a precise guide to establishing,

[2] In more complicated decision problems, there may be more options, or the options may depend on more than just the truth of $A$. We focus here on the simplest case, set out in this section, for two reasons. Firstly because already the simplest case forces some interesting constraints on Bayesianism, as we will see in subsequent sections. Second, as is demonstrated in Appendix 1, the analysis presented here is robust under a natural complexification of the decision problem, in which establishing $\bar{A}$ is also included as an extra option.

[3] If the sum of the false advantage and the true advantage is zero, then the threshold is not well defined and the decision-theoretic account offers no guidance. For ease of exposition, we set this case aside in what follows.

and any violation of this decision rule would violate the norms of Bayesianism. Any more detailed Bayesian analysis of establishing must validate this decision rule, on pain of inconsistency.

## §4
## Subjective degrees of belief?

Most philosophical proponents of Bayesianism advocate a subjective version of Bayesianism, according to which it is rationally permissible to adopt any degree of belief in a contingent proposition $A$, in the absence of evidence that bears on the truth of $A$. Interestingly, the Bayesian account of establishing rules out subjective Bayesianism, as we shall see in this section.

In addition to norms on decision making, Bayesianism offers a guide to setting initial degrees of belief and to changing degrees of belief in the light of new evidence. With respect to the former task, norms can be divided into three kinds. Firstly, there is some structural norm, which provides constraints on the structure of degrees of belief. Standardly, this says that degrees of belief should be probabilities—representable by a probability function $P_E$ defined over the domain of expressible propositions. Second, there are also evidential norms, which ensure that rational degrees of belief are appropriately constrained by evidence. For example: $P_E(A) = 1$ if $E$ logically implies $A$. Most Bayesians also maintain that degrees of belief should be calibrated to chances where available: $P_E(A) = x$ if $E$ says that the chance of $A$ is $x$, as long as there is no other evidence in $E$ that defeats this ascription of degree of belief; we will refer to this sort of evidential norm as a calibration norm. Another sort of evidential norm (a deference norm) requires, or at least permits, deference to an appropriate expert in the absence of other relevant evidence: $P_E(A) = x$ if $E$ says that some rational expert believes $A$ to degree $x$, as long as there is no other evidence in $E$ more pertinent to $A$. Third, there are equivocation norms, which ensure that rational degrees of belief are equivocal in the absence of evidence, but which are not endorsed by all Bayesians. For example: $P_E(A) = 0$ or $1$ only if forced by evidence or by the axioms of probability. Another sort of equivocation norm requires that $P_E(A) = 1/2$ if there is no evidence that bears on the truth of $A$, where $A$ is a contingent atomic proposition (this is a version of the principle of indifference). This norm is often generalised to require that $P_E$ be the probability function with maximum entropy from all those that satisfy structural and evidential norms (the maximum entropy principle). The adoption of an equivocation norm is a feature that distinguishes objective Bayesianism from subjective Bayesianism: these norms can substantially limit the scope for subjective choice with respect to initial degrees of belief.

Although subjective Bayesianism stops short of equivocation norms, it does impose strong constraints on changes to degrees of belief. The standard approach here is to take $P_E(\cdot)$ to be a conditional probability function, $P(\cdot|E)$. Then, on new evidence $F$ that is compatible with $E$, subjective Bayesians argue that your degree of belief in $A$ should change from $P_E(A) = P(A|E)$ to $P_{EF}(A) = P(A|EF)$ (Bayesian conditionalisation). Thus the 'prior' probability function $P(\cdot|\cdot)$ fully determines all 'posterior' degrees of belief. Some objective Bayesians advocate Bayesian conditionalisation (e.g., Jaynes, 2003) while others reject the identification of conditional beliefs with conditional probabilities and argue that degrees of belief should be updated by reapplying the maximum entropy principle to the new evidence (Williamson, 2010). We will not need to decide between these approaches to updating here. We shall just note that on the former approach, the norm that requires calibration of degrees of belief to chances is called the 'principal principle'.

Having introduced the key norms of Bayesian epistemology, we shall now see that the Bayesian account of establishing requires some equivocation norm. Consider a simple example. Suppose $A$ says that Angie has *Angiostrongylus* infection, that the threshold for establishing $A$, $\tau_A$, is 0.9, and that there is no evidence available that bears on the truth of $A$. In the absence of such evidence, subjective Bayesianism would deem any degree of belief in $A$ to be rationally permissible. In particular, $P_E(A) = 0.91$ is rationally permissible. But if you do adopt this degree of belief then $P_E(A) > \tau_A$ and you ought to establish $A$ and treat it as evidence for other propositions, such as the proposition that Angie has meningitis; establishing $A$ would also motivate particular treatment decisions. Clearly, however, it is not rationally permissible to take $A$ as established in the total absence of evidence that bears on the truth of $A$: this is the platitude of non-vacuity, introduced in §2. Since subjectivism fails to validate this platitude, it fails to offer a viable account of establishing. Some normative constraint is required to ensure that $P_E(A) < 0.9$, that is, to ensure that this degree of belief is sufficiently equivocal in the absence of evidence that bears on the truth of $A$. This is an objective Bayesian constraint.

One might object to this argument as follows: here we cannot assume a total absence of evidence relevant to $A$, because the Bayesian decision-theoretic account of establishing presupposes a utility matrix and these utilities tell us something about preferences relating to $A$, so they provide evidence relevant to $A$. However, this objection misses the mark: the above argument does not assume a total absence of information pertaining to $A$, but rather an absence of evidence that bears on the truth of $A$. The utilities of establishing and of not establishing $A$ do not provide evidence that $A$ is true, nor evidence that $A$ is false. Therefore, these utilities do not on their own provide grounds for establishing $A$. Grounds for establishing $A$ must be grounds for taking $A$ to be true: they must provide evidence *for* $A$, not merely evidence that concerns $A$.[4]

## §5
## Subjective utilities?

So far, we have seen that a Bayesian account of establishing requires objective constraints on degrees of belief that force non-extreme values in the absence of evidence. This goes against the dominant philosophical view of Bayesianism, which is subjectivism. Another tenet that Bayesians hold dear is that utilities are subjective: although there are normative structural constraints on preferences or utilities, the preference relations or utility values themselves are personal, with any utilities that satisfy the structural norms being rationally permissible (see, e.g., Savage, 1954).

[4]These considerations point to another potential disanalogy between believing (in the qualitative sense, rather than the quantitative, Bayesian sense) and establishing. Some suggest that utilities alone can provide grounds for qualitative belief—e.g., Pascal put forward his famous wager to argue for believing that God exists, purely on the basis of a comparison of the utilities of believing and of not believing. In contrast, it is not plausible to suggest that utilities alone can provide grounds for establishing. In particular, one cannot rationally establish that God exists in the absence of evidence that bears on the truth of whether God exists. That this is so can be seen by appealing to the connection between the act of establishing and the corresponding evidential relation: to establish a proposition is to imply that the proposition is true, so far as one can tell from the evidence. Since evidence that does not bear on the truth of God's existence does not establish God's existence, it is rationally impermissible to establish God's existence without evidence that bears on the truth of God's existence.

We will see in this section that the Bayesian account of establishing challenges this tenet too.

Consider a possible subjective utility matrix (relating, as usual, to a contingent atomic proposition $A$):

| | $A$ | $\bar{A}$ |
|---|---|---|
| establish $A$ | 10 | −3 |
| don't establish $A$ | −10 | 2 |

Here, the false advantage is 5 and the true advantage is 20, so the threshold for establishing is 5/25, $\tau_A = 0.2$. Suppose again that you have no evidence that bears on the truth of $A$. In such a case, any Bayesian—subjective or objective—will deem $P_E(A) = 0.5$ to be rationally permissible. But such a value passes the threshold for establishing, so, if you do have this degree of belief, a Bayesian account of establishing requires that you establish $A$, adding it to your evidence and fully believing it thereafter. This cannot be right: by non-vacuity, it is not rationally permissible to establish a proposition in the total absence of evidence that bears on the truth of that proposition. Thus the above utility matrix is pathological, leading to irrationality, and must be deemed rationally impermissible.

In order to avoid this problem, we clearly need a higher threshold for establishing: the threshold will need to be greater than $\frac{1}{2}$, at the very least. Now, $\tau_A > \frac{1}{2}$ just when $s_{\bar{A}} - e_{\bar{A}} > s_A - e_A$, i.e., just when

$$\text{false advantage} > \text{true advantage}.$$

To be viable, then, a Bayesian account of establishing needs to impose this normative constraint on utilities.

This is perhaps surprising, because there does not seem to be anything wrong with the above utility matrix from a subjective point of view: it simply represents the utilities of someone who sees more advantage in establishing $A$ when it is true than in not establishing $A$ when it is false. Indeed, it is not the Bayesian connection between degrees of belief and decisions that renders this utility matrix impermissible. Its impermissibility only comes to light when the act of establishing enters the mix: Bayesianism concerns evidence as well as degree of belief and decision making, and it turns out that this further normative constraint is required if Bayesianism is to accommodate the non-vacuity of establishing.

Although it might be surprising that there is the need for such a normative constraint, the constraint itself is descriptively quite plausible. First, it is plausible that $e_{\bar{A}}$ is much less than $e_A$: it is normally much worse to take a false proposition as evidence than not to take a true proposition as evidence. Taking a false proposition as evidence removes it from contention and enables it to be used as evidence for other propositions, with the potential to lead to many more erroneous inferences. On the other hand, failing to take a true proposition as evidence has only a mild impact on the intellectual economy: some time and effort may be spent in keeping it under scrutiny and evaluating its probability, and one may miss out on some sound inferences while it is being evaluated. Second, it is often plausible that $s_{\bar{A}}$ is of similar magnitude to $s_A$. Establishing $A$ when true is good for the intellectual economy because resources need no longer be devoted to its evaluation and because it offers opportunities for new sound inferences. On the other hand, keeping $A$ under scrutiny when it is false is good because it leaves open the possibility that $\bar{A}$ might be established in the future, and in the meantime, inferences will be based on proper assessment of its probability. If, indeed, $e_{\bar{A}}$ is much less than $e_A$ but $s_{\bar{A}}$ is of similar magnitude to $s_A$, then the false advantage will exceed the true advantage, as the normative constraint on utilities requires.

The upshot is that a Bayesian account of establishing requires non-structural constraints on utilities. One can provide a general argument for this conclusion as follows. Suppose for reductio that there are no non-structural constraints on utilities. Consider a case in which there is no evidence that bears on the truth of $A$. Since by assumption there are no non-structural constraints on utilities, it is rationally permissible to have a utility matrix for establishing $A$ that is the same as the utility matrix for $\bar{A}$, and in which the false advantage is no greater than the true advantage. In which case, the threshold $\tau_A$ for establishing $A$ and the threshold $\tau_{\bar{A}}$ for establishing $\bar{A}$ are both no greater than $\frac{1}{2}$. Now, one of $P(A)$ and $P(\bar{A}) = 1 - P(A)$ must be $\frac{1}{2}$ or greater. Hence it is deemed permissible to establish at least one of $A, \bar{A}$. But this contradicts non-vacuity. Therefore, there must be non-structural constraints on utilities after all.[5]

Suppose, then, that the false advantage is some multiple of the true advantage, $s_{\bar{A}} - e_{\bar{A}} = x(s_A - e_A)$ where $x > 1$. Then

$$\tau_A = \frac{x}{x+1}.$$

The following table provides examples of thresholds for establishing that correspond to various possible values of $x$:

| $x$ | $\tau_A$ |
|---|---|
| 1.5 | 0.6 |
| 2 | 2/3 |
| 3 | 0.75 |
| 9 | 0.9 |
| 19 | 0.95 |
| 99 | 0.99 |

## §6
## Stability and deference

### §6.1. The deference problem

We have seen that both probabilities and utilities require objective constraints if Bayesianism is to provide a viable account of the act of establishing. In this section we will investigate a rather different question. We observed in §2 that establishing requires stability of belief in addition to threshold belief. On the Bayesian account, however, the decision to establish $A$ seems to depend only on whether the degree of belief in $A$ meets a threshold ($\tau_A$, defined in §3). The question then arises as to whether the Bayesian account of establishing can adequately accommodate the need for stability of degree of belief.

Consider an example. Suppose that nothing in evidence $E$ bears on the truth of whether Angie has *Angiostrongylus* infection (proposition $A$), and your threshold for establishing $A$ is 2/3. Suppose, next, proposition $Y$: a doctor says that she believes $A$ to degree 0.67. Let us assume for the moment that it is rationally permissible to

[5] Appendix 2 shows that this argument also goes through when establishing $\bar{A}$ is included alongside establishing $A$ as an explicit option in the decision problem.

defer to the expert and calibrate your own degree of belief to hers, and you do so, setting $P_{YE}(A) = 0.67$. The Bayesian account then says that you ought to establish $A$. This is problematic if your degree of belief in $A$ is likely to be volatile in the light of further evidence.

Your degree of belief in $A$ might be volatile where the doctor's evidence, although more extensive than your own, is nevertheless somewhat limited—e.g., where new test results are due which are likely to lead to substantial changes to her degree of belief, and thus to your own degree of belief should you continue to defer to her. Alternatively, it may be the case that the doctor's expertise, although greater than your own, is nevertheless rather limited, and that you will shortly defer to a consultant who is more of a specialist in the area and who will be providing an opinion in due course. Either way, if it is likely that your degree of belief will change, then there is a significant chance that it will soon fall below the threshold for establishing $A$.

In such a scenario, the Bayesian account has it that you are in a position now to establish $A$ but there is a substantial chance that shortly you should not take $A$ to be established. This is problematic because it conflicts with the stability condition for establishing: when you establish a proposition you add it to your stock of evidence and, while establishing is fallible, you would not expect to revisit that proposition and revoke its status as evidence, at least in the short term. To say 'we have established the existence of a new particle but we are not confident that we will take it to be established tomorrow' is little short of Moore-paradoxical.[6] The ideal that, once established, a proposition remains in the evidence base is a fundamental Bayesian presupposition, underlying the use of Bayesian conditionalisation to update degrees of belief in the light of new evidence, for example. In order to conditionalise, the evidence base can only grow—Bayesian conditionalisation says nothing about what to do if some item of evidence is revoked; it only offers a guide to life when evidence is stable.

This problem for the Bayesian account of establishing arose because we assumed that it is rationally permissible to defer unequivocally to the degrees of belief of an expert, even when those degrees of belief are based on limited evidence or where the expertise itself is limited (as is invariably the case in practice). To avoid the problem, then, we must avoid this assumption, in the hope of developing a version of Bayesianism in which stability of degree of belief is required for a degree belief to surpass a threshold for establishing. Only if such a version of Bayesianism is possible will the Bayesian account of establishing capture the requirement of stability of degree of belief in addition to that of threshold degree of belief.

Thus, Bayesianism needs what might be called 'discounted deference': a principled way of ensuring that it is not normally permissible to defer fully to an expert's degree of belief, but that it is nevertheless rational to be influenced to some extent by that degree of belief. What is needed is a compromise between a wholesale adoption of the doctor's degree of belief, 0.67, and being fully equivocal, represented by degree of belief 0.5, which one might adopt if there is no evidence at all that bears on the truth of $A$ (in particular, if there is no evidence of the base rate of *Angiostrongylus* infection). The need for this compromise could be motivated by the thought that one should not defer fully to anything but chance.

[6] This paradox is not the preface paradox: it may be reasonable to say that we are confident that *some* proposition we have established will be revoked in due course, even though it would not be reasonable to say of any particular established proposition that we have substantial confidence that it will be revoked.

Bayesians who identify conditional beliefs with conditional probabilities might attempt to accommodate discounted deference by applying the theorem of total probability. Where $Y$ says that the expert believes $A$ to degree 0.67 and $X_x$ says that the chance of $A$ is $x$,

$$P(A|YE) = \int_x P(A|YX_xE)P(X_x|YE)\, dx$$

Now, the truth of $A$ is already determined, so its chance is 0 or 1. Hence,

$$\begin{aligned} P(A|YE) &= 0 \cdot P(X_0|YE) + 1 \cdot P(X_1|YE) \\ &= P(X_1|YE) \\ &= P(A|YE) \end{aligned}$$

So the theorem of total probability offers no substantive constraint in this case.

The theorem of total probability offers no help even if $A$ is a proposition whose chance is not 0 or 1. In this case,

$$\begin{aligned} P(A|YE) &= \int_x P(A|YX_xE)P(X_x|YE)\, dx \\ &= \int_x xP(X_x|YE)\, dx \end{aligned}$$

because the principal principle implies that $P(A|YX_xE) = x$. Thus to apply the theorem of total probability here, one needs to specify the values $P(X_x|YE)$ for all $x \in [0,1]$. This distribution can be thought of as a representation of the perceived reliability of the expert. There are two problems with this proposal. First, specifying all these values is a big ask of the deliberating agent. If Bayesianism is to be a practical guide to life, it would be better not to have to specify all these values. Second, there is no guarantee that this proposal will lead to discounted deference. If there is no evidence of the expert's reliability, the probabilities $P(X_x|YE)$ are not constrained by any evidential norm. According to subjective Bayesianism, any values for these probabilities are rationally permissible. Thus on a subjective Bayesian account there is no guarantee that $P(A|YE)$ takes a value between 0.5 and 0.67. Without such a guarantee, non-vacuity would still be violated: it would be deemed rationally permissible to establish $A$ in the absence of evidence that bears on the truth of $A$. On an objective Bayesian account, on the other hand, one would need to specify some equivocal distribution on the unit interval to specify $P(X_x|YE)$. Now, determining an equivocal distribution over a continuous domain is problematic at the best of times (see, e.g., Keynes, 1921), but in this case the task is harder still because that equivocal distribution needs to embody some default assumption of reliability of the expert. Just specifying a uniform distribution on the unit interval will yield a value of 0.5 for $P(A|YE)$, which fails to accommodate the fact that it is rationally permissible to be influenced to some positive degree by a competent expert. Arguably, one should give more credence to the claim that the chance of $A$ is within an interval around 0.67 than to the claim that it is in an interval of the same size around 0.33, say, on the basis of $Y$. But if, as is plausible, the reliability distribution $P(X_x|YE)$ is symmetric around 0.67, then $P(A|YE)$ will take the value 0.67, which constitutes wholesale deference rather than discounted deference. Therefore, on either a subjective or an objective version of Bayesianism, an appeal to the theorem of total probability fails to offer an account of discounted deference that overcomes the problem of stability.

There is, however, another option for developing an account of discounted deference. The approach taken here is to treat the problem of discounted deference as analogous to the problem of determining inductive probabilities and to modify a solution to the inductive problem that was developed in Williamson (2013) and Williamson (2017, Chapter 7) so that it can solve the problem of discounted deference. First, in §6.2, I introduce the inductive problem and the proposed solution; then in §6.3 I apply this solution to the deference problem.

### §6.2. Determining inductive probabilities

The problem of determining inductive probabilities can be understood as follows. A calibration norm such as the principal principle says that one should defer wholesale to chances, where these chances are available. In practice, however, we only have imperfect evidence of chances, usually provided by taking a finite sample and finding a proportion in that sample. Thus, to estimate the chance that Angie has *Angiostrongylus* infection, one might consider the proportion of patients that have *Angiostrongylus* infection in a sample of patients whose relevant characteristics are similar to Angie's. Unless that sample is very large, one should not defer wholesale to the sample proportion. If the sample proportion is 0.67, for example, it would be appropriate to adopt a degree of belief in $A$ that is somewhere between 0.5 and 0.67, in absence of any other evidence that bears on the truth of $A$. The inductive problem is the problem of determining exactly which value or values in this range are appropriate.

The larger the sample size, the greater the influence of the sample proportion. Perhaps at some point, the sample is so large that one is prepared to grant that the sample proportion is an accurate estimate of the chance and fully defer to it via the calibration norm. But in other situations, one is not prepared to grant this. In these other situations, one can use an interval estimate instead of a point estimate: one might be prepared to grant that the chance is within some interval around the sample proportion, even though one is not prepared to grant that the chance is equal to the sample proportion. Confidence interval estimation methods are standardly used to determine an interval estimate of a chance. The confidence level will depend on the benefit of correctly estimating the chance and the cost of incorrectly estimating the chance; I argue in Williamson (2017, Chapter 7) that Bayesian decision theory can in principle be applied to choose the confidence level. Note that the confidence level can be, and usually is, chosen in advance of collecting the sample: one makes a commitment to grant that the chance is in the confidence interval determined by the sample and the chosen confidence level, and that confidence interval is the smallest interval for which one is prepared to grant this. Thus one commits to a confidence level, and subsequently the sample yields a sample proportion and one takes the chance to be within the corresponding confidence interval.

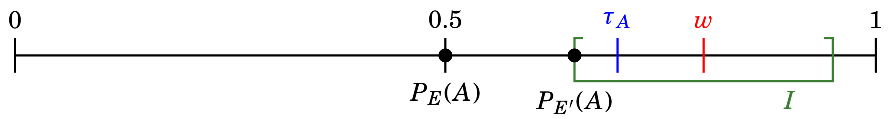
Suppose, then, that $W_w$ says that a sample proportion pertinent to $A$ is $w$ and suppose that other evidence $E$ does not bear on $A$. Let $I$ be the smallest closed interval for which you are prepared to grant that the chance is in that interval.[7] After sampling you thus establish that the chance of $A$ is in interval $I$; we denote this established proposition by $X_I$. Your new evidence is $E' = X_I W_w E$. The principal principle then forces your new degree of belief to lie in the interval $I$, $P_{E'}(A) \in I$.

[7]There is always some such interval, because you should be prepared to grant that the chance is at least within the unit interval.

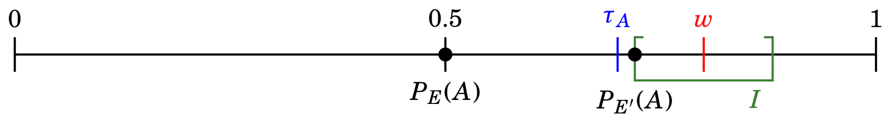
This can be seen as follows:

$$\begin{aligned} P(A|X_I W_w E) &= \int_{x\in[0,1]} P(A|X_I W_w X_x E) P(X_x|X_I W_w E)\, dx \\ &= \int_{x\in I} x P(X_x|X_I W_w E)\, dx \end{aligned}$$

which is in $I$ since $P(X_x|X_I W_w E) \in [0,1]$ and $\int_{x\in I} P(X_x|X_I W_w E)\, dx = 1$. Finally, if, as argued in §4, an equivocation norm is in force, then $P_{E'}(A)$ must take some sufficiently equivocal value within the interval $I$, that is, a value in $I$ near 0.5. (The maximum entropy principle, for example, would, in the absence of other evidence, choose the value in $I$ that is closest to 0.5.) This procedure is depicted as follows:




Here $P_E(A) = 0.5$, since, it is supposed, $E$ contains no evidence that bears on the truth of $A$. $E'$ constrains your degree of belief in $A$ to lie in the interval $I$ around the sample proportion $w$, and the most equivocal value—the lower end of the interval—is the only appropriate value for $P_{E'}(A)$ according to the maximum entropy principle.[8] Only if $P_{E'}(A) > \tau_A$ will it be appropriate to establish $A$. This requires two things: that $w > \tau_A$ and that $I$ is small enough that the lower end of the interval $I$ exceeds the threshold $\tau_A$. This second condition will be met only if the confidence interval is sufficiently small:




Since it is established that the chance of $A$ is in the interval $I$, future degrees of belief in $A$ can be expected to remain in this interval: in that sense, they are stable. Thus we see that, in the inductive case, $A$ is established only if the threshold is met and there is sufficient stability of degrees of belief.

### §6.3. A deference principle

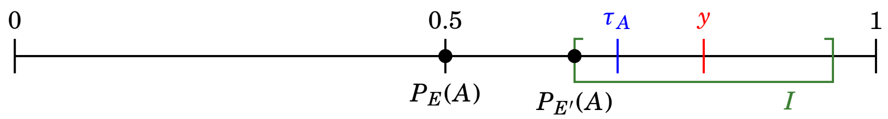
Having seen how the stability requirement can be accommodated in the context of the problem of determining inductive probabilities, we shall now see that this kind of approach can be applied to the analogous problem of discounted deference. This will allow us to alleviate the concern about stability raised in §6.1.

First we consider a principle governing wholesale deference that is analogous to a calibration norm such as the principal principle. Let $Z_z$ say that a fully competent

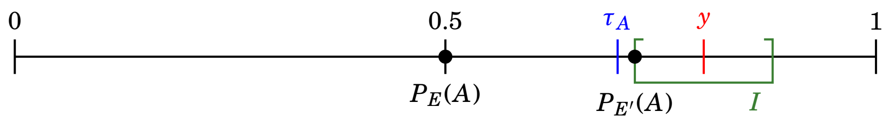
[8] It is important to emphasise that $P_{E'}(A)$ is a rational degree of belief, *not* an estimate of the chance of $A$. $w$ remains the best point estimate of the chance of $A$, but since this point estimate is a real number, one should not give any positive credence to the proposition that the chance of $A$ coincides precisely with $w$. If all one is prepared to grant is that the chance is in the interval $I$, then the principal principle *only* requires that one's degree of belief should lie within that interval. The reasons for choosing an equivocal value (a value close to 0.5) from within that interval are non-evidential (Williamson, 2018).

and rational expert, with ideal evidence of $A$, ought to believe $A$ to some degree $z$. Then, a deference principle that is analogous to a calibration norm would force $P_{Z_z E}(A) = z$, as long as there is no evidence in $E$ that overrides this ascription of degree of belief. Here, 'ideal evidence' can be taken to consist of all the evidence one would collect if the cost of evidence gathering were no object. (Ideal evidence cannot be construed as all matters of fact up to the present time, because these facts would include $A$ or $\bar{A}$ and the deference principle would trivialise.)

In practice, one is presented not with $Z_z$, but $Y_y$, the proposition that the actual expert's degree of belief is $y$. Nevertheless, as in the case of calibration to chance, one might be prepared to infer something about the ideal value $z$ from the actual expert's degree of belief $y$. Consider the smallest closed interval $I$ for which one is rational to establish $Z_I$ on the basis of $Y_y E$. (Trivially, one is at least rational to establish $Z_{[0,1]}$ on the basis of $Y_y E$.) Letting $E' = Z_I Y_y E$ and applying the idealised deference principle, we have that $P_{E'}(A) \in I$, just as in the analogous case of the principal principle, considered above. Applying an equivocation norm, $P_{E'}(A)$ will be a value in $I$ near 0.5—the value closest to 0.5, according to the maximum entropy principle. This yields an account of discounted deference that is closely analogous to the account of induction presented above:




Only if $P_{E'}(A) > \tau_A$ will it be appropriate to establish $A$:




This requires both that $y > \tau_A$ and that $I$ is sufficiently small. A small interval $I$ corresponds to a narrow range of values for the degree of belief $z$ of a fully competent and rational expert with ideal evidence of $A$. Since this ideal degree of belief is the value to which one's own degree of belief should defer and it is established that $z$ is in the narrow interval $I$, the expectation is that one's own degrees of belief will remain in this small interval. This is an expectation of stability of degree of belief. Thus this approach overcomes the problem posed in §6.1: both stability and threshold are features of establishing.

Two questions face this new account of discounted deference. First, how do we calculate $I$? In the case of determining inductive probabilities, confidence interval methods were natural, but here there is no obvious off-the-shelf statistical method. One possible approach is to think of the doctor's expertise as distilling experience of previous patients—both her own experience and the experience passed on to her through her training—and to think of her credence as a sample proportion based on this sample of previous patients. If one can quantify the expertise of the doctor by identifying some number $n$ such that one would view a sample of size $n$ as equally informative about $A$ as the expert's degree of belief, one can then exploit the analogy between discounted deference and the problem of determining inductive probabilities and one can use confidence interval methods to choose $I$.

As before, one might apply expected utility theory to determine a confidence level, and one can determine a confidence interval by thinking of the agent's credence $y$ as a proportion generated by a sample of size $n$.[9]

The second question that arises is: how does this new account of discounted deference interact with a calibration norm such as the principal principle? One might think that the principal principle should always trump deference to expert belief, because chances, where one can obtain them, offer the best evidence one could possibly hope for. Perhaps. However, we cannot normally obtain chances—only estimates of chances. If we are indifferent between an expert's opinion and a proportion from a sample of size $n$, then, clearly, a sample proportion $w$ generated by a sample of size $n$ should not simply trump the expert's credence $y$, because they are viewed as equally informative about $A$. In this case, we might think of the evidence as containing two samples of size $n$ and use statistical methods for combining studies, such as meta-analysis, to generate a confidence interval based on all the available evidence.

## §7
## Conclusions and objections

To sum up, Bayesianism *can* provide a viable account of establishing, but only a particular kind of Bayesianism can. In order to provide a viable account of establishing, one needs to accept non-evidential constraints on probabilities and non-structural constraints on utilities, and one needs to accept that it is normally impermissible to defer wholesale to an expert's degrees of belief: one needs a satisfactory account of discounted deference. One cannot get such an account of discounted deference by applying the principle of total probability, but one can by viewing deference as analogous the problem of determining inductive probabilities and by adapting the approach of Williamson (2017). This yields a very idiosyncratic version of Bayesianism: a version of objective Bayesianism that combines frequentist and Bayesian methods and that is prescriptive about utilities. The challenge to advocates of other versions of Bayesianism is: how can your versions yield viable accounts of establishing?

We close by considering some possible objections to this account of establishing.

One might argue that establishing is not really an act and that Bayesian decision theory is therefore the wrong tool to apply. The thought here might be that it is evidence that establishes a claim, not an agent. This objection is hard to mesh with the practice of science, since establishing propositions appears to be the primary activity of scientists. If establishing were not an act, there would seem to be very little for the scientist to do. This objection is also hard to reconcile with law, medicine and other fields where establishing is a core activity, if not the primary activity. In the account developed in §6.3, the agent is central is central to establishing: utilities are to some extent agent-relative, and hence so are the confidence levels; moreover, judgements need to be made about the equivalence between an expert's opinion and a sample. In scientific contexts, there is room for debate about these judgements and thus room to settle disagreements, but these judgements are nonetheless attributable to the deliberating agent. Establishing is very much an act.

[9] One complication here is that $y$ is the expert's degree of belief, not her estimate of the chance of $A$, and so $y$ may be more equivocal than the expert's best estimate of chance, particularly when the expert has little or no evidence to go on. Thus this approach is most applicable where the expert has plenty of relevance experience and evidence.

A related worry is that establishing may not be an intentional act: perhaps establishing is a subconscious act and we have no control over what we establish, in which case Bayesian decision theory may yet be the wrong tool to apply. Two considerations help to alleviate this worry. First, while some acts of establishing may be performed subconsciously, not all are. To take a concrete example, the International Agency for Research on Cancer assesses whether certain chemicals and other agents are carcinogenic: at review meetings, one subgroup of reviewers explicitly decides whether they can establish from epidemiological studies that the agent causes cancer in humans, while another subgroup decides whether they can establish from animal studies that the agent causes cancer in experimental animals (IARC, 2019, pp. 31–32). These explicit decision problems are common in science and medicine, and Bayesian decision theory is applicable to such situations. Second, even in cases of unintentional establishing, it can be helpful to consciously re-evaluate these cases. While a hallucination may subconsciously cause one to establish the presence of giant spiders and to treat this proposition as evidence that one should run away, it will be helpful to consciously re-evaluate this proposition in due course. When the context changes, Bayesianism can offer a guide to life even with respect to those propositions that, at first, one cannot help but strongly believe or establish.

Another concern is that establishing is prone to standard problems that face analyses of propositional attitudes in terms of threshold degrees of belief. For example, it is sometimes suggested that a qualitative belief can be understood as a degree of belief that meets an appropriate threshold. Such accounts are prone to lottery and preface paradoxes, for instance. In the case of establishing, it may be that one ought to establish $A$ and one ought to establish $B$ but one ought not to establish $AB$, if $P_E(A) > \tau_A$ and $P_E(B) > \tau_B$ but $P_E(AB) < \tau_{AB}$. This appears incongruous, not least because if one does establish $A$ and one does establish $B$ then it will usually be at least permissible to establish $AB$, since $P_{EAB}(AB) = 1$.

A similarly odd phenomenon arises with a lottery example. Suppose $E$ specifies a fair $n$-ticket lottery with a prize of 1000 som and a ticket cost of 1 som, and $A_i$ says that ticket $i$ loses. Suppose also that establishing $A_i$ would lead you not to buy ticket $A_i$, but that you commit to buy the smallest-numbered ticket that has not been established to lose, if there are any such tickets. Initially (before any tickets have been established to lose) the utility matrix for establishing $A_1$ might look like this:

| | $A_1$ | $\bar{A}_1$ |
|---|---|---|
| establish $A_1$ | 0 | 0 |
| don't establish $A_1$ | $-1$ | 999 |

Here, if ticket 1 loses ($A_1$) then establishing that it loses will result in no loss or gain, because it will not be purchased. Similarly, if it wins ($\bar{A}_1$) then establishing that it loses will result in no gain or loss. If it is not established to lose, then you will buy that ticket, in which case you will pay 1 and have an overall gain of 999 if that ticket wins. The threshold for establishing $A_1$ is thus $\tau_{A_1} = 999/1000$. Equivocating, your credence that ticket 1 wins is $1/n$, where $n$ is the number of tickets in the lottery, so $P_E(A_1) = (n-1)/n$. According to this analysis, you should establish $A_1$ if there are more than a thousand tickets, $n > 1000$.

Suppose you do indeed establish $A_1$. Consider $A_2$ next. The utility matrix is the same, but now $P_{EA_1}(A_2) = (n-2)/(n-1)$. In general, $P_{EA_1,\ldots,A_k}(A_{k+1}) = (n-$

$k-1)/(n-k)$. There will come a point, then, when $(n-k-1)/(n-k) < 999/1000$; at this point it will not be permissible to establish $A_{k+1}$. For example, if $n = 2000$ then one should establish that the first 999 tickets lose, it is permissible to establish that ticket 1000 will lose, but if one does also establish $A_{1000}$ then it is not permissible to establish $A_{1001}$. This is a peculiar phenomenon, given that each ticket has exactly the same probability of winning.

While this is not the place for a general discussion of lottery and related paradoxes, which affect many theories in formal epistemology, it is worth noting that the account of establishing developed here has one line of response that is not available to many other theories. To the extent that these phenomena are problematic consequences of decisions to establish, they have negative utility. It is possible to take this negative utility into account when deciding whether to establish a proposition. Thus the utility matrices in the lottery example might be better represented as follows:

| | $A_i$ | $\bar{A}_i$ |
|---|---|---|
| establish $A_i$ | $-u$ | $-u$ |
| don't establish $A_i$ | $-1$ | 999 |

If $u \geq 1$ then $\tau_{A_i} \geq 1$ and for no $i$ is it the case that one is rationally required to establish $A_i$. This avoids incurring the negative utility engendered by establishing that some lottery tickets will lose, but not others.

**Acknowledgements**

This research was supported by funding from the Leverhulme Trust (grant RPG-2019-059) and the Deutsche Forschungsgemeinschaft (DFG, grant LA 4093/3-1). I am very grateful to David Corfield, Mikayla Kelley and the anonymous referees for helpful comments.

## Appendix 1: Robustness of the analysis of §3

In this appendix we see that the lessons of the simple decision scenario of §3 also hold of the more complex decision problem that arises when we include establishing $\bar{A}$ as an option.

We abbreviate the acts *establish* $A$ and *don't establish* $A$ by ✓$A$ and ✗$A$ respectively. The decision table for establishing $A$ (from §3) is:

| | $A$ | $\bar{A}$ |
|---|---|---|
| ✓$A$ | $s_A$ | $e_{\bar{A}}$ |
| ✗$A$ | $e_A$ | $s_{\bar{A}}$ |

We saw that this decision table requires establishing $A$ when

$$P(A) > \frac{\text{false advantage}}{\text{true advantage} + \text{false advantage}}$$

The analogous decision table for establishing $\bar{A}$ is:

| | $A$ | $\bar{A}$ |
|---|---|---|
| ✓$\bar{A}$ | $e'_A$ | $s'_{\bar{A}}$ |
| ✗$\bar{A}$ | $s'_A$ | $e'_{\bar{A}}$ |

By reasoning analogous to that of §3, in this modified decision problem one ought to establish $\bar{A}$ just when

$$P(A) < \frac{\text{true}' \text{ advantage}}{\text{true}' \text{ advantage} + \text{false}' \text{ advantage}}$$

where the true′ advantage is $s'_{\bar{A}} - e'_{\bar{A}}$ and the false′ advantage is $s'_A - e'_A$.

Now consider a more complex decision scenario in which one might establish one or both of $A, \bar{A}$. Arguably, one can represent this scenario as follows:

| | $A$ | $\bar{A}$ |
|---|---|---|
| ✓$A$ ✓$\bar{A}$ | $-u$ | $-u$ |
| ✓$A$ ✗$\bar{A}$ | $s_A + s'_{\bar{A}}$ | $e_{\bar{A}} + e'_{\bar{A}}$ |
| ✗$A$ ✓$\bar{A}$ | $e_A + e'_{\bar{A}}$ | $s_{\bar{A}} + s'_{\bar{A}}$ |
| ✗$A$ ✗$\bar{A}$ | $e_A + s'_{\bar{A}}$ | $s_{\bar{A}} + e'_{\bar{A}}$ |

One can often take utilities to be additive. However, if both $A$ and $\bar{A}$ are established then evidence becomes inconsistent which has very negative utility, denoted by $-u$ in this table, where $u$ is very large. The expected utility of ✓$A$ ✓$\bar{A}$ is then $-u$, which, we may suppose, rules this option out of contention.

Given that this option is ruled out of contention, establishing $A$ is rationally required when the expected utility of ✓$A$ ✗$\bar{A}$ exceeds that of ✗$A$ ✗$\bar{A}$. It is not hard to see that this is the case just when

$$P(A) > \frac{\text{false advantage}}{\text{true advantage} + \text{false advantage}}$$

That this concurs with the criterion derived from the first decision matrix set out above shows that the analysis of §3 is robust under a natural complexification of the decision problem.

Moreover, according to the modified decision scenario, establishing $\bar{A}$ is rationally required when the when the expected utility of ✗$A$ ✓$\bar{A}$ exceeds that of ✗$A$ ✗$\bar{A}$. It is not hard to see that this is the case just when

$$P(A) < \frac{\text{true}' \text{ advantage}}{\text{true}' \text{ advantage} + \text{false}' \text{ advantage}}$$

Again, this provides evidence of robustness of the analysis. Note that the same points can be made about the rational permissibility of establishing $A$ or $\bar{A}$ if we change the strict inequalities above to non-strict inequalities.

### Appendix 2: Robustness of the analysis of §5

In this appendix we see that the general argument for the need for non-structural constraints on utilities presented in §5 is robust under moving to the more complex decision problem of Appendix 1.

Let us rerun the general argument in the context of the complex decision problem of Appendix 1, with the four options ✓$A$ ✓$\bar{A}$, ✓$A$ ✗$\bar{A}$, ✗$A$ ✓$\bar{A}$, ✗$A$ ✗$\bar{A}$.

Suppose for reductio that there are no non-structural constraints on utilities, and consider a case in which there is no evidence that bears on the truth of $A$. By assumption there are no non-structural constraints on utilities, so it is rationally permissible to have a utility matrix for establishing $A$ that is the same as the utility matrix for $\bar{A}$. In the notation of Appendix 1,

$$s'_A = s_{\bar{A}}, e'_A = e_{\bar{A}}, s'_{\bar{A}} = s_A, e'_{\bar{A}} = e_A.$$

These constraints imply that:

$$\text{true}' \text{ advantage} = \text{true advantage}$$

$$\text{false}' \text{ advantage} = \text{false advantage}$$

By assumption it is also rationally permissible that the false advantage is no greater than the true advantage:

$$\text{false advantage} \leq \text{true advantage}$$

This latter inequality implies that:

$$\frac{\text{false advantage}}{\text{true advantage} + \text{false advantage}} \leq 1/2 \leq \frac{\text{true advantage}}{\text{true advantage} + \text{false advantage}}$$

From the analysis of Appendix 1, we see then that establishing $A$ is rationally permissible if $P(A) \geq 1/2$ and establishing $\bar{A}$ is rationally permissible if $P(A) \leq 1/2$. So whatever value $P(A)$ takes, it is rationally permissible to establish at least one of $A$ and $\bar{A}$. But this contradicts non-vacuity. Thus there must be non-structural constraints on utilities after all.

## Bibliography


Butler, B. (1736). *The analogy of religion.* J.M. Dent & Sons, London, 1906 edition.

IARC (2019). *IARC Monographs on the Evaluation of Carcinogenic Hazards to Humans: Preamble.* Lyon. https://monographs.iarc.fr/wp-content/uploads/2019/01/Preamble-2019.pdf.

Jaynes, E. T. (2003). *Probability theory: the logic of science.* Cambridge University Press, Cambridge.

Jeffrey, R. (2004). *Subjective probability: the real thing.* Cambridge University Press, Cambridge.

Kaplan, M. (1981). A Bayesian theory of rational acceptance. *The Journal of Philosophy*, 78(6):305–330.

Keynes, J. M. (1921). *A treatise on probability.* Macmillan, London, 1973 edition.

Leitgeb, H. (2017). *The stability of belief: how rational belief coheres with probability.* Oxford University Press, Oxford.

Mitova, V. (2017). *Believable evidence.* Cambridge University Press, Cambridge.

Parkkinen, V.-P., Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M. P., Norell, C., Russo, F., Shaw, B., and Williamson, J. (2018). *Evaluating evidence of mechanisms in medicine: principles and procedures.* Springer, Cham, Switzerland.

Rowbottom, D. P. (2014). Information versus knowledge in confirmation theory. *Logique et Analyse*, 57(226):137–149.

Savage, L. J. (1954). *The Foundations of Statistics.* Dover, New York, second (1972) edition.

Skyrms, B. (1977). Resiliency, propensities, and causal necessity. *The Journal of Philosophy*, 74(11):704–713.

Williamson, J. (2010). *In defence of objective Bayesianism.* Oxford University Press, Oxford.

Williamson, J. (2013). Why frequentists and Bayesians need each other. *Erkenntnis*, 78(2):293–318.

Williamson, J. (2015). Deliberation, judgement and the nature of evidence. *Economics and Philosophy*, 31(1):27–65.

Williamson, J. (2017). *Lectures on inductive logic.* Oxford University Press, Oxford.

Williamson, J. (2018). Justifying the principle of indifference. *European Journal for the Philosophy of Science*, 8:559–586.

Williamson, T. (2000). *Knowledge and its limits.* Oxford University Press, Oxford.