



Pettigrew, R. (2016). Jamesian epistemology formalised: An explication of 'The Will to Believe'. *Episteme*, 13(3), 253-268. <https://doi.org/10.1017/epi.2015.44>

Peer reviewed version

Link to published version (if available):
[10.1017/epi.2015.44](https://doi.org/10.1017/epi.2015.44)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via Cambridge University Press at <https://doi.org/10.1017/epi.2015.44> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Jamesian epistemology formalised: an explication of ‘The Will to Believe’

Famously, William James held that there are two commandments that govern our epistemic life.

There are two ways of looking at our duty in the matter of opinion, — ways entirely different, and yet ways about whose difference the theory of knowledge seems hitherto to have shown very little concern. *We must know the truth*; and *we must avoid error*, — these are our first and great commandments as would be knowers; but they are not two ways of stating an identical commandment [...] Believe truth! Shun error! — these, we see, are two materially different laws; and by choosing between them we may end by coloring differently our whole intellectual life. We may regard the chase for truth as paramount, and the avoidance of error as secondary; or we may, on the other hand, treat the avoidance of error as more imperative, and let truth take its chance. (James, 1896, Section VII)

In this paper, I give a formal account of James’ claim using the tools of epistemic utility theory. I begin by giving the account for categorical doxastic states — that is, full belief, full disbelief, and suspension of judgment. Then I will show how the account plays out for graded doxastic states — that is, credences. The latter part of the paper thus answers a question left open in (Pettigrew, 2014). Konek (ta) gives a related treatment of imprecise credences.

It is not entirely clear whether James intends, in ‘The Will to Believe’, to speak of beliefs and disbeliefs or of credences. He certainly talks of “options” between “hypotheses”, which suggests the choice between two categorical states — belief in one hypothesis or belief in the other. But he also talks of different strengths of a “believing tendency” and suggests that only a hypothesis with the “maximum of liveness” (presumably the maximum “believing tendency”) counts as a belief (James, 1896, Section I). In any case, in this paper, we treat both.

1 Epistemic utility theory

According to epistemic utility theory, the rationality of a doxastic state is determined by how conducive it is to obtaining epistemic utility. Thus, like utilitarianism in ethics and other forms of consequentialism, it makes the good prior to the right; a doxastic state is epistemically right (or rational) if it conduces to what is epistemically good (to wit, epistemic utility). Investigations in this area thus comprise two parts: first, we give an account of the epistemic utility of the doxastic states in which we are interested; second, we state some consequentialist principles, drawn from decision theory, that govern choices between different options that are assigned different utilities depending on how the world is.

For instance, take Jim Joyce’s non-pragmatic vindication of the credal norm of probabilism (Joyce, 1998). First, he gives an account of epistemic utility: credences have greater epistemic utility the greater their accuracy. Second, he appeals to a decision-theoretic principle: it is the principle of dominance, which says that if one option is better than another in all situations — in the jargon, the former dominates the latter — then the latter is irrational. Finally, he derives probabilism by showing that any credences that violate that norm are accuracy dominated.

In this paper, we’ll attempt to reconstruct — perhaps better, explicate — William James’ claim concerning the two commandments of epistemology using the same notion of epistemic utility to which Joyce appeals. But we’ll be interested in a rather different decision rule from the dominance principle that features in Joyce’s argument.

2 Categorical doxastic states

In this part of the paper, we are concerned with epistemic norms that govern categorical doxastic states. I take there to be three types of such states: full belief, full disbelief, and suspension of judgement. Thus, we model an agent’s categorical doxastic states at a given time by her belief function b , which takes each proposition X that she entertains and returns B if she believes X , D if she disbelieves X , and S if she suspends judgment on X . Let \mathcal{F} be the set of propositions that the agent entertains. Then $b : \mathcal{F} \rightarrow \{B, D, S\}$.

In James’ terminology, \mathcal{F} is the set of “live hypotheses”. And the agent then has, for each proposition X in \mathcal{F} , a “forced choice” between believing, disbelieving, and suspending on X . James writes: “to say [...] ‘Do not decide but leave the question open,’ is itself a passional decision, just like deciding ‘yes’ or ‘no,’ and is attended with the same risk of losing the truth” (James, 1896, 334).

Our question is this: Which belief functions are rational for which agents? To answer it using epistemic utility theory, we first need to answer two subsidiary questions: How should we measure the epistemic utility of a belief function? What decision principles should guide our choice of belief function in the presence of that measure of epistemic utility? Recall

William James' two commandments — *Believe truth! Shun error!* How we weigh the relative importance of one against the other determines how we manage our epistemic life. In fact, it turns out that the answers to both of our subsidiary questions — the question of which measure of epistemic utility to use and the question of which decision principle to adopt — will be affected by the relative importance we ascribe to James' two imperatives.

2.1 Jamesian measures of epistemic utility

First, the measure of epistemic utility. In fact, in this paper, as in Joyce's, it will be most convenient to talk about measures of epistemic *disutility*. But such measures are easily obtained from measures of epistemic utility: the negative of an epistemic utility function is an epistemic disutility function; and vice versa. James' two commandments suggest a *veritistic* or *accuracy-based* account of epistemic utility. That is, it seems reasonable to interpret James as taking the sole fundamental source of epistemic utility to be the having of accurate belief states — *Believe truth!* — and the not having of inaccurate ones — *Shun error!*. This suggests that the epistemic disutility of a belief function b at a world w is determined as follows.¹ First, there is a *local inaccuracy measure* $\mathfrak{s} : \{0, 1\} \times \{B, D, S\} \rightarrow [0, \infty]$. The idea is this: $\mathfrak{s}(1, B)$ gives the inaccuracy of having a belief in a true proposition, while $\mathfrak{s}(0, B)$ gives the inaccuracy of having a belief in a false proposition. Similarly, $\mathfrak{s}(1, S)$ gives the inaccuracy of suspending judgment in a proposition that's true, and $\mathfrak{s}(0, S)$ gives the inaccuracy of suspending in a proposition that's false. And again $\mathfrak{s}(1, D)$ gives the inaccuracy of disbelieving a truth, while $\mathfrak{s}(0, D)$ gives the inaccuracy of disbelieving a falsehood. We make the following assumptions:

- $\mathfrak{s}(1, B) = \mathfrak{s}(0, D)$. That is, a false disbelief is as inaccurate as a true belief. We denote this $-R$, where $R > 0$. 'R' for getting it right.
- $\mathfrak{s}(1, D) = \mathfrak{s}(0, B)$. That is, a false belief is as inaccurate as a true disbelief. We denote this W , where $W > 0$. 'W' for getting it wrong.
- $\mathfrak{s}(1, S) = \mathfrak{s}(0, S) = 0$. That is, a suspension of judgment always has inaccuracy 0, regardless of the truth or falsity of the proposition.

Now we can use our local inaccuracy measure \mathfrak{s} to define a *global inaccuracy measure* \mathfrak{I} for categorical doxastic states. This takes an entire belief function b defined on \mathcal{F} and a possible world w and returns a measure of the inaccuracy of b at w :

$$\mathfrak{I}(b, w) = \sum_{X \in \mathcal{F}} \mathfrak{s}(v_w(X), b(X))$$

¹Many consequences of measuring the epistemic utility of beliefs and disbeliefs in this way have been explored by Hempel (1962), Levi (1967), Easwaran (ms), Easwaran and Fitelson (ta), and Fitelson (ms). This section complements that work by explaining the consequences of a risk-sensitive decision principle in the presence of this account of epistemic utility.

where $v_w(X) = 1$ if X is true and $v_w(X) = 0$ if X is false. That is, the inaccuracy of a belief function is simply the sum of the inaccuracies of the individual categorical attitudes that comprise it. Thus, a belief function that assigns belief (B) to X and disbelief (D) to \bar{X} , evaluated at a world at which X is true will have global inaccuracy $\mathfrak{s}(1, B) + \mathfrak{s}(0, D) = -2R$.

Thus, a global inaccuracy measure is determined by a local inaccuracy measure; and a local inaccuracy measure is determined by the badness we assign to a true belief/false disbelief — that is, the value of R — and the badness we assign to a true disbelief/false belief — that is, the value of W . It is in this latter endeavour that we might think the balance between James' commandments makes a difference. If we “regard the chase for truth as paramount, and the avoidance of error as secondary”, then we will take R to be greater than W ; if, on the other hand, we “treat the avoidance of error as more imperative, and let truth take its chance”, then we will take W to exceed R . And of course it's also possible to weight them equally and let $W = R$. Depending on what we choose in this case, quite different epistemic norms follow from decision-theoretic principles.

For instance, consider the dominance principle to which Joyce appeals in his argument for probabilism. And consider an agent who entertains only two propositions, X and its negation \bar{X} . If $R \geq W$, then the belief functions that are not accuracy dominated are these: believe X and disbelieve \bar{X} ; believe \bar{X} and disbelieve X ; believe X and believe \bar{X} ; disbelieve X and disbelieve \bar{X} . That is, if we weigh *Believe truth!* more highly than *Shun error!*, then believing each of a pair of mutually inconsistent propositions is not ruled out as irrational by the dominance principle (nor is disbelieving both).² On the other hand, if $W > R$, then the belief functions that are not accuracy dominated are these: believe X and disbelieve \bar{X} ; believe \bar{X} and disbelieve X ; suspend judgment on X and \bar{X} . That is, if we weigh *Shun error!* more highly than *Believe truth!*, then believing each of a pair of mutually inconsistent propositions is ruled out as irrational by the dominance principle (as is disbelieving each of them).

2.2 Jamesian decision principles

This gives us a taste of the consequences of different weightings of James' two imperatives when it comes to measuring the epistemic disutility of a categorical doxastic state. But these weightings do not only affect the way in which we measure epistemic disutility. They also affect the decision principle that we use in conjunction with that measure of epistemic disutility to assess the rationality of categorical doxastic states. After all, James' two commandments each encode opposite attitudes to epistemic risk. To believe or disbelieve a proposition gives you a shot at maximal epistemic utility (that is, minimal inaccuracy), however you set R and W : you will receive maximal epistemic utility if what you believe turns out to be true, or if what you disbelieve turns out to be false. But of course it also opens you up to the pos-

²Due to this counterintuitive consequence of $R \geq W$, Easwaran and Fitelson (ta) simply assume that $W > R$.

sibility of minimal epistemic utility (that is, maximal inaccuracy): you will receive minimal epistemic utility if what you believe turns out to be false, or what you disbelieve turns out to be true. Suspension, on the other hand, does not give you a shot at maximal epistemic utility, but nor does it open you up to the possibility of minimal epistemic utility. Thus, believing and disbelieving are risky doxastic attitudes in a way that suspending is not.

In ‘The Will to Believe’, James seems more concerned with this reading of his competing commandments than with the reading explicated in §2.1 above. Thus, he talks of suspending judgement as being “attended with the [...] risk of losing the truth” (James, 1896, 334) and the “awful risk of believing lies” (James, 1896, 338). However, at other points, where he also uses the language of risk, he is clearly imagining that this will be spelled out in terms of epistemic utility, such as when he writes: “worse things than being duped may happen to a man in this world” (James, 1896, 339).

We know from practical decision theory that there are two ways that we might encode an agent’s sensitivity to risk. We might try to build some of the agent’s attitudes to risk into her utility function; or we might capture it in the decision principle that she endorses (Buchak, 2014). For instance, in practical decision theory, we might represent a risk-averse agent as assigning diminishing marginal utility to money: that is, we might say that her utility for an outcome is a concave function of the money she will receive in that outcome. But we might also take her to endorse a risk-averse decision principle. Such principles include: Minimax, which says that an agent should act to minimise her worst-case disutility; Minimax Regret, which says that an agent should act to minimise her worst-case regret; and certain versions of the Hurwicz Criterion, which takes the weighted sum of the best-case disutility of an option and its worst-case disutility and says that an agent should minimise that. In section 2.1, we saw how we might encode an agent’s attitudes to epistemic risk into her epistemic utility function: a risk-averse agent, who weighs *Believe truth!* more heavily than *Shun error!*, will assign a greater penalty to being wrong — believing a falsehood or disbelieving a truth — than she will assign a reward to being right — believing a truth or disbelieving a falsehood. A risk-seeking agent, who weighs the commandments in the opposite way, will do the opposite. In this section, we are interested in encoding those attitudes in a decision principle. The natural choice is the Hurwicz Criterion, since that allows us to weigh two competing attitudes to risk differently; and, as we will see below, those two attitudes correspond to James’ two commandments.

Let us begin by stating the Hurwicz Criterion as a decision principle for general decision theory — that is, decision theory in which we make no assumptions about the nature of the options between which the agent is choosing. We will then adapt it to give a decision principle for epistemic utility theory, and we will describe the epistemic norm that we derive from it in conjunction with our account of epistemic utility from section 2.1.

Suppose \mathcal{O} is a set of options and \mathcal{W} is the set of possible worlds. Now suppose that \mathfrak{U} is

a utility function that takes each option and each possible world and returns the utility of the outcome of choosing that option in that world. Thus $-\mathfrak{U}$ is a disutility function. Then, as I said above, the Hurwicz Criterion asks an agent to minimise a weighted sum of the best-case disutility of an option and its worst-case disutility. Thus, different versions of the Hurwicz Criterion are given by different choices of weight. Let $0 \leq \lambda \leq 1$. Then, for any option o in \mathcal{O} , define

$$H_\lambda^{-\mathfrak{U}}(o) := \lambda \min_{w \in \mathcal{W}} -\mathfrak{U}(o, w) + (1 - \lambda) \max_{w \in \mathcal{W}} -\mathfrak{U}(o, w)$$

So λ is the weight assigned to the best-case scenario; the scenario in which the disutility of the option is minimal. Then:

Hurwicz $_\lambda$ Criterion Suppose $o, o^* \in \mathcal{O}$. $H_\lambda^{-\mathfrak{U}}(o^*) < H_\lambda^{-\mathfrak{U}}(o) \Rightarrow o$ is irrational for an agent with utility function \mathfrak{U} .

That is, all options o for which $H_\lambda^{-\mathfrak{U}}(o)$ is not minimal are ruled irrational. Thus, suppose our agent is faced with the following decision problem, where the numbers in the table represent the disutility of the relevant outcomes:

	<i>Rain</i>	<i>No Rain</i>
<i>Umbrella</i>	10	6
<i>No Umbrella</i>	12	0

Then

- $H_\lambda(\text{Umbrella}) = 6\lambda + 10(1 - \lambda) = 10 - 4\lambda$
- $H_\lambda(\text{No Umbrella}) = 0\lambda + 12(1 - \lambda) = 12 - 12\lambda$

Thus, for $\lambda < \frac{1}{4}$, *Umbrella* is the rational option; for $\lambda = \frac{1}{4}$, both are rationally permitted; and for $\lambda > \frac{1}{4}$, *No Umbrella* is the rational option. Thus, as λ increases and more weight is given to the best-case situation, the agent becomes more risk-seeking and less risk-averse.

Now suppose that \mathfrak{J} is a global inaccuracy measure — that is, an epistemic disutility function for belief functions. Then, for a belief function b ,

$$H_\lambda^{\mathfrak{J}}(b) := \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) + (1 - \lambda) \max_{w \in \mathcal{W}} \mathfrak{J}(b, w)$$

Again, λ weights the best-case (or minimal) disutility. Then the Hurwicz $_\lambda$ Criterion demands:

Hurwicz $_\lambda$ Criterion (categorical; epistemic) Suppose b and b^* are belief functions on \mathcal{F} . $H_\lambda^{\mathfrak{J}}(b^*) < H_\lambda^{\mathfrak{J}}(b) \Rightarrow b$ is irrational for an agent who measures inaccuracy using \mathfrak{J} .

Thus, λ represents the weight given to the Jamesian commandment *Believe truth!*, since that is the risk-seeking commandment, which enjoins us to try to be accurate even though doing so inevitably exposes us to the risk of being inaccurate. On the other hand, $1 - \lambda$ represents the weight given to *Shun error!*, since that is the risk-averse commandment, which enjoins us to avoid exposing ourselves to the risk of inaccuracy, even if it is only by doing so that we open ourselves to the possibility of accuracy.

Now, suppose \mathfrak{J} is determined by the local inaccuracy measure \mathfrak{s} , and that \mathfrak{s} is in turn determined by R and W . Then our question is this: What epistemic norm follows from Hurwicz_λ Criterion when it is applied to \mathfrak{J} ? To state the norm, we first define what it means to be a *Hurwicz $_\lambda$ belief function*. It turns out that these belief functions are precisely those permitted by **Hurwicz $_\lambda$ Criterion (categorical; epistemic)**.

Suppose we are considering belief functions defined on an algebra of propositions \mathcal{F} . Our first job is to say when a probability function defined on \mathcal{F} *probabilifies* a belief function on \mathcal{F} relative to R and W . There are three cases that we must consider: $W > R$, $W = R$, $W < R$. Let's take them in turn:

Definition 1 (p probabilifies b) Let p be a probability function defined on \mathcal{F} .

(i) Suppose $W > R$. Then p probabilifies b if, for each X in \mathcal{F} , one of the following holds:

- $\frac{W}{R+W} < p(X)$ and $b(X) = B$
- $\frac{W}{R+W} = p(X)$ and $b(X) = B$ or $b(X) = S$
- $\frac{R}{R+W} < p(X) < \frac{W}{R+W}$ and $b(X) = S$
- $p(X) = \frac{R}{R+W}$ and $b(X) = D$ or $b(X) = S$
- $p(X) < \frac{R}{R+W}$ and $b(X) = D$.

Thus, roughly speaking, p probabilifies b if p and b jointly satisfy the Lockean thesis for belief with respect to the belief threshold $\frac{W}{R+W}$. That thesis, again roughly, says that an agent believes X if she assigns sufficiently high probability to X ; she disbelieves X if she assigns sufficiently low probability to X ; and otherwise she suspends. On the threshold itself, she has some choice.

Thus, for instance, if \mathcal{F} contains only God exists and God does not exist, and $R = 1$ and $W = 2$, then p probabilifies a belief in God exists and disbelief in God does not exist if $p(\text{God exists}) \geq \frac{2}{3} = \frac{2}{2+1}$.

(ii) Suppose $W = R$. Then p probabilifies b if, for each X in \mathcal{F} , one of the following holds:

- $\frac{1}{2} < p(X)$ and $b(X) = B$
- $\frac{1}{2} = p(X)$ and $b(X) = S$ or $b(X) = D$ or $b(X) = B$
- $p(X) < \frac{1}{2}$ and $b(X) = D$

(iii) Suppose $W < R$. Then p probabilifies b if, for each X in \mathcal{F} , one of the following holds:

- $\frac{1}{2} < p(X)$ and $b(X) = B$
- $\frac{1}{2} = p(X)$ and $b(X) = D$ or $b(X) = B$
- $p(X) < \frac{1}{2}$ and $b(X) = D$

The crucial result is the following, due to Hempel (1962) and Easwaran (ms):

Theorem 1 p probabilifies $b \Leftrightarrow b$ has minimal expected epistemic disutility by the lights of p .

That is, p probabilifies $b \Leftrightarrow$

$$\text{Exp}_{\mathfrak{J}}(b|p) \leq \text{Exp}_{\mathfrak{J}}(b'|p)$$

for all belief functions b' , where $\text{Exp}_{\mathfrak{J}}(b|p) := \sum_{w \in \mathcal{W}} p(w)\mathfrak{J}(b, w)$.

That's the first stage in our definition of the Hurwicz_{λ} belief functions. In the second stage, we will define them as the belief functions that are probabilified by a particular sort of probability function.

Suppose \mathcal{F} is a finite algebra. Then let $\mathcal{W}_{\mathcal{F}} = \{w_1, \dots, w_n\}$ be the atoms of that algebra: one might think of these as the maximally specific possibilities relative to \mathcal{F} . Now, for each $1 \leq k \leq n$, we need to define the k^{th} Hurwicz_{λ} probability function on \mathcal{F} . We denote it p_k^{λ} and we define it as follows:

Definition 2 (p_k^{λ})

(i) Suppose $\frac{1}{n} \leq \lambda$. Then

$$p_k^{\lambda}(w_j) := \begin{cases} \lambda & \text{if } j = k \\ \frac{1-\lambda}{n-1} & \text{if } j \neq k \end{cases}$$

(ii) Suppose $\lambda < \frac{1}{n}$. Then

$$p_k^{\lambda}(w_j) := \frac{1}{n}$$

Thus, suppose you are thinking about the colour of my newest tie. You partition the space of possibilities into three: red, yellow, blue. Thus, you entertain only the propositions in the algebra generated by the following three possibilities: $w_1 = \text{Red}$, $w_2 = \text{Blue}$, $w_3 = \text{Yellow}$. And suppose $\lambda = \frac{1}{2}$. Then $p_1^{\lambda}(\text{Red}) = \lambda = \frac{1}{2}$, while $p_1^{\lambda}(\text{Blue}) = \frac{1-\lambda}{n-1} = \frac{1}{4}$ and $p_1^{\lambda}(\text{Yellow}) = \frac{1-\lambda}{n-1} = \frac{1}{4}$. If, on the other hand, $\lambda = \frac{1}{3}$, then $p_1^{\lambda}(\text{Red}) = p_1^{\lambda}(\text{Blue}) = p_1^{\lambda}(\text{Yellow}) = \frac{1}{3}$.

As we will see, these probability functions play an equally important role when we turn to ask what the consequences of the Hurwicz_{λ} Criterion are in the case of credences. Here is the crucial theorem that states the consequences of the Hurwicz_{λ} Criterion in the case of categorical doxastic states:

Theorem 2 If b satisfies Hurwicz_{λ} Criterion (categorical; epistemic), then b is probabilified by p_k^{λ} for some $1 \leq k \leq n$.

The proof is given in the Appendix. Let's see this result in action. I ask you to pick a number from 1 to 9. Let X_k be the proposition that says that you pick k . Thus, I entertain the propositions X_1, \dots, X_9 .³ Now suppose that $R = 1$ and $W = 3$. Thus, when I pick my epistemic disutility function, I assign three times greater weight to James' commandment to shun error than to his commandment to believe truth. Now let us see what happens when we change the relative weights that I assign to these commandments when I pick the version of the Hurwicz Criterion that I will apply in the presence of that epistemic disutility function.

First, note that, for any $\frac{1}{9} < \lambda \leq 1$, $p_k^\lambda(X_j) = \frac{1-\lambda}{8} < \frac{1}{4} = \frac{R}{R+W}$ (for $j \neq k$). Thus, since **Hurwicz $_\lambda$ Criterion (categorical; epistemic)** demands that I have belief function that is probabilified by p_k^λ for some $1 \leq k \leq 9$, and since any such belief function will assign disbelief to eight out of the nine propositions X_1, \dots, X_9 , I must disbelieve eight out of the nine propositions. What **Hurwicz $_\lambda$ Criterion (categorical; epistemic)** demands in the case of the remaining proposition depends on λ . If $\frac{3}{4} < \lambda$, then it demands that I believe the remaining proposition; if $\lambda = \frac{3}{4}$, it permits me to believe or suspend; if $\frac{1}{4} \leq \lambda < \frac{3}{4}$, it demands that I suspend; and if $\lambda < \frac{1}{4}$, it demands that I disbelieve.

Second, note that, if $\lambda \leq \frac{1}{9}$, then $p_k^\lambda(X_j) = \frac{1}{9}$ for all $1 \leq j \leq 9$. Thus, since $\frac{1}{9} < \frac{1}{4}$, p_k^λ probabilifies a belief function that assigns disbelief to each of the nine propositions. Of course, this gives rise to a situation akin to the Lottery Paradox, where we disbelieve each of a set of mutually exclusive and exhaustive propositions. Recall that we saw above that, if $R > W$, there were inconsistent sets of propositions such that believing each is not accuracy dominated. The implications of the **Hurwicz $_\lambda$ Criterion (categorical; epistemic)** just sketched show that such attitudes may be permitted even if $R < W$.

3 Graded doxastic states

In this section, we turn from categorical to graded doxastic states. Thus, instead of representing an agent by her belief function, we represent her by her credence function. This is the function that takes each proposition that the agent entertains — that is, James' "live hypotheses" — and returns her credence in that proposition, where, by convention, we measure credence on a continuous scale from the minimal credence of 0 to the maximal credence of 1. Thus, $c : \mathcal{F} \rightarrow [0, 1]$.

We measure the inaccuracy of a credence function in much the same way that we measure the inaccuracy of a belief function. Thus, we begin by defining a local inaccuracy measure $\mathfrak{s} : \{0, 1\} \times [0, 1] \rightarrow [0, \infty]$. Thus, for $0 \leq x \leq 1$, $\mathfrak{s}(1, x)$ measures the inaccuracy of having credence x in a true proposition, while $\mathfrak{s}(0, x)$ measures the inaccuracy of having credence x in a false proposition. And again we define the global inaccuracy measure \mathfrak{J} by summing

³If we wished to make this more Jamesian, we might let X_1, \dots, X_9 be eight different theistic hypotheses — e.g., theosophist, unitarian — and the atheist hypothesis. But it is easier to see what is going on if we keep the example simple.

the local inaccuracies given by \mathfrak{s} :

$$\mathfrak{I}(c, w) = \sum_{X \in \mathcal{F}} \mathfrak{s}(v_w(X), c(X))$$

3.1 Jamesian measures of epistemic disutility

We are now in a position to ask how our attitudes to the Jamesian commandments might be reflected in the way that we measure the inaccuracy of credence functions.⁴ Recall that, in the case of belief functions, we said that the agent who weighs *Believe truth!* more heavily than *Shun error!* will assign a greater reward to getting it right than she will assign a penalty to getting it wrong. Thus, for such agents — whom we might think of as epistemic risk-seekers — moving from the neutral position of suspension to a belief introduces the possibility of penalty and reward, but the potential reward has more goodness than the penalty has badness. And the opposite is true of agents — whom we might think of as epistemic risk-avoiders — who weighs *Shun error!* more heavily than *Believe truth!*. For them, moving from the neutral position of suspension introduces the possibility of penalty and reward, but the potential reward has less goodness than the penalty has badness. Translating this to the credal case, we get the following: an epistemic risk-seeker should have a local inaccuracy measure where a move away from neutral ground — which we might take to be the point at which $\mathfrak{s}(1, x) = \mathfrak{s}(x, 0)$ — will introduce the possibility of penalty and reward, but the potential reward — that is, the decrease in inaccuracy if the move is in the right direction — will have more goodness than the penalty — that is the increase in inaccuracy if the move is in the wrong direction — will have badness. And the opposite for the epistemic risk-avoider. More precisely:

- If \mathfrak{s} is a *risk-seeking local inaccuracy measure* and x^* is the neutral point for \mathfrak{s} — that is, $\mathfrak{s}(1, x^*) = \mathfrak{s}(0, x^*)$ — then, for all x and ε :

- If $x > x^*$ and $x + \varepsilon \leq 1$, then

$$\mathfrak{s}(1, x) - \mathfrak{s}(1, x + \varepsilon) > \mathfrak{s}(0, x + \varepsilon) - \mathfrak{s}(0, x)$$

- If $x < x^*$ and $x - \varepsilon \geq 0$, then

$$\mathfrak{s}(0, x) - \mathfrak{s}(0, x - \varepsilon) > \mathfrak{s}(1, x - \varepsilon) - \mathfrak{s}(1, x)$$

- If \mathfrak{s} is a *risk-avoiding local inaccuracy measure* and x^* is the neutral point for \mathfrak{s} — that is, $\mathfrak{s}(1, x^*) = \mathfrak{s}(0, x^*)$ — then, for all x and ε :

⁴See (Joyce, 2009, Section 9) for a related account of how James' commandments might affect choice of credal inaccuracy measure.

– If $x > x^*$ and $x + \varepsilon \leq 1$, then

$$\mathfrak{s}(1, x) - \mathfrak{s}(1, x + \varepsilon) < \mathfrak{s}(0, x + \varepsilon) - \mathfrak{s}(0, x)$$

– If $x < x^*$ and $x - \varepsilon \geq 0$, then

$$\mathfrak{s}(0, x) - \mathfrak{s}(0, x - \varepsilon) < \mathfrak{s}(1, x - \varepsilon) - \mathfrak{s}(1, x)$$

Here is an example of a risk-avoiding local inaccuracy measure. It is called the *quadratic scoring rule* and it is defined as follows: $\mathfrak{q}(1, x) = (1 - x)^2$ and $\mathfrak{q}(0, x) = x^2$. And here is an example of a risk-seeking local inaccuracy measure. It is called the *square root scoring rule* and it is defined as follows: $\mathfrak{r}(1, x) = \sqrt{1 - x}$ and $\mathfrak{r}(0, x) = \sqrt{x}$.

Recall from above: If $R < W$ — that is, the agent has a risk-averse epistemic utility function — then only the consistent categorical doxastic states are not dominated, but if $R > W$ — that is, the agent’s epistemic utility function is risk-seeking — then inconsistent states are undominated. Similarly, it turns out, the credence functions that are undominated relative to the inaccuracy measure generated by the quadratic scoring rule are precisely the coherent ones — that is, the ones that satisfy the axioms of the probability calculus — whereas there are incoherent credence functions that are undominated relative to the inaccuracy measure generated by the square root scoring rule. The former fact is a theorem due to de Finetti (1974) and generalised by Joyce (1998), Predd et al. (2009), and Pettigrew (ms); the latter is easily seen by considering the credence function that assigns 0 to X and 0 to \bar{X} .

3.2 Jamesian decision principles

Finally, we turn to the question of decision principles that reflect different attitudes to the Jamesian commandments. In fact, we have very little work to do here, since the natural decision principle was introduced above. It is the Hurwicz Criterion. Here is the version that pertains to credences. Given a global inaccuracy measure \mathfrak{J} for credence functions — that is, an epistemic disutility function for those states — let:

$$H_\lambda^{\mathfrak{J}}(c) := \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(c, w) + (1 - \lambda) \max_{w \in \mathcal{W}} \mathfrak{J}(c, w)$$

Hurwicz $_\lambda$ Criterion (graded; epistemic) Suppose c and c^* are credence functions on \mathcal{F} . $H_\lambda^{\mathfrak{J}}(c^*) < H_\lambda^{\mathfrak{J}}(c) \Rightarrow c$ is irrational for an agent who measures inaccuracy using \mathfrak{J} .

So our question is: Which credence functions are ruled irrational by the **Hurwicz $_\lambda$ Criterion (graded; epistemic)** and which are permitted? Suppose the credence functions between which we are choosing are all defined on a finite algebra \mathcal{F} with $\mathcal{W}_{\mathcal{F}} = \{w_1, \dots, w_n\}$, as above. Then, under a natural assumption, it turns out to be only the Hurwicz $_\lambda$ probability

functions on \mathcal{F} , defined above, that are permitted (not ruled irrational) by **Hurwicz $_{\lambda}$ Criterion (graded; epistemic)**. The natural assumption is this: our local inaccuracy measure is strictly proper.

Definition 3 A local inaccuracy measure \mathfrak{s} is strictly proper if, for all $0 \leq p \leq 1$,

$$p\mathfrak{s}(1, x) + (1 - p)\mathfrak{s}(0, x)$$

is uniquely minimised as a function of x at $x = p$.

That is, \mathfrak{s} is strictly proper if a given credence p in a proposition X will expect itself to be the least inaccurate amongst all possible credences in X . A number of epistemic utility theorists have made this assumption or something similar (Joyce, 2009; Greaves and Wallace, 2006; Pettigrew, ms). However, it is worth noting that while the quadratic scoring rule q defined above is strictly proper, the other scoring rule τ is not. Nonetheless, if we make this assumption, we can identify the credence functions permitted by **Hurwicz $_{\lambda}$ Criterion (graded; epistemic)** by appealing to the following theorem:

Theorem 3 Suppose \mathfrak{s} is strictly proper. Then, if c satisfies **Hurwicz $_{\lambda}$ Criterion (graded; epistemic)**, then $c = p_k^{\lambda}$ for some $1 \leq k \leq n$.

Thus, if we consider again the situation in which I ask you to pick a number from 1 to 9, and proposition X_i says that you pick number i , then we have the following. If I am very risk averse and assign $\lambda \leq \frac{1}{9}$, then my credence should be simply the uniform distribution over the nine possibilities: $c_k^{\lambda}(X_k) = \frac{1}{9}$. But, on the other hand, if I am more willing to take epistemic risks, and I give more weight to the best-case disutility by setting $\frac{1}{9} < \lambda$, then I should plump for one of the possible numbers k (with $1 \leq k \leq 9$) and have credence λ that you picked k , while dividing my credences equally over the other eight options, thus assigning $\frac{1-\lambda}{8}$ to each.

4 Conclusion

This concludes our attempt to give a formal account of James' two commandments and the different weightings that we might assign to them. Where does it leave us? James' purpose in drawing attention to these two commandments was to explain how it might be rational to go beyond one's evidence. He wished to rebut W. K. Clifford's famous claim that "it is wrong always, everywhere, and for every one to believe anything upon insufficient evidence" (Clifford, 1877). In the absence of evidence favouring one hypothesis over any other, we might think we are obliged to assign equal credence to each, or to suspend judgement on each, as demanded by the Principle of Indifference or the Principle of Insufficient Reason. James wished to show that in fact this is not rationally required. For an agent who weighs *Believe*

truth! more heavily than *Shun error!*, it can be rational to pick a hypothesis from amongst the set between which our evidence fails to distinguish and assign a doxastic pro-attitude — beliefs, for instance, or high credences — to that while maintaining neutrality or doxastic anti-attitudes — suspensions or disbeliefs, for instance, or middle to low credences — to the rest. Our investigation has vindicated James. The formal results presented above show that, for an agent who endorses risk-seeking versions of the epistemic Hurwicz Criterion, or risk-seeking epistemic utility functions, exactly such doxastic states are required.

This conclusion has consequences for the debate between theists and atheists, as James anticipated. But it also has consequences for the debate over scepticism about the external world. Indeed, it has consequences for any debate with the following structure: there is a range of mutually exclusive and exhaustive hypotheses; the evidence does not tell in favour of any one over any other; and yet we wish to say that having a strong epistemic attitude in favour of one of those hypotheses — a belief, for instance, or a high credence — is rationally permissible. For instance, we wish to say that it is rationally permissible to deny the sceptical hypothesis and believe that the external world exists, even though the Cartesian demon argument shows that all the evidence we could ever acquire could not decide between them. And similarly for debates over scepticism about other minds, about the efficacy of induction, and so on. These judgments of rational permissibility are vindicated by James' account, which we have here made precise. For an epistemic risk-seeker, it is rational to plump for one hypothesis amongst the range of those available and believe that.

5 Appendix: proofs of Theorems 2 and 3

In this section, I present the proof of Theorem 2. The proof of Theorem 3 is exactly analogous.

We prove Theorem 2 by proving its claim for each of two cases, labelled CASE I and CASE II below. In CASE I, we consider the case in which $\frac{1}{n} < \lambda$. In CASE II, we consider the case in which $\lambda \leq \frac{1}{n}$.

CASE I First, let us assume that $\frac{1}{n} < \lambda$. We prove the following lemma:

Lemma 4 Suppose $\frac{1}{n} < \lambda$.

(I) If b is probabilified by p_k^λ , then

$$\text{Exp}_{\mathfrak{J}}(b|p_k^\lambda) = H_\lambda^{\mathfrak{J}}(b)$$

(II) If b is not probabilified by p_j^λ for any $1 \leq j \leq n$, and $\mathfrak{J}(b, w_k) \leq \mathfrak{J}(b, w_j)$ for all $1 \leq j \leq n$, then

$$\text{Exp}_{\mathfrak{J}}(b|p_k^\lambda) \leq H_\lambda^{\mathfrak{J}}(b)$$

Proof of Lemma 4(I). Note that, if b is probabilified by p_k^λ , then b achieves minimal inaccuracy at world w_k . Thus:

$$\begin{aligned}
& \text{Exp}_{\mathfrak{J}}(b|p_k^\lambda) \\
&= \sum_{i=1}^n p_k^\lambda(w_i) \mathfrak{J}(b, w_i) \\
&= \frac{1-\lambda}{n-1} \mathfrak{J}(b, w_1) + \dots + \lambda \mathfrak{J}(b, w_k) + \dots + \frac{1-\lambda}{n-1} \mathfrak{J}(b, w_n) \\
&= \frac{1-\lambda}{n-1} \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) + \dots + \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) + \dots + \frac{1-\lambda}{n-1} \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) \\
&= (1-\lambda) \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) + \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) \\
&= H_\lambda^{\mathfrak{J}}(b)
\end{aligned}$$

Proof of Lemma 4(II). By assumption, b achieves minimal inaccuracy at world w_k . So:

$$\begin{aligned}
& \text{Exp}_{\mathfrak{J}}(b|p_k^\lambda) \\
&= \sum_{i=1}^n p_k^\lambda(w_i) \mathfrak{J}(b, w_i) \\
&= \frac{1-\lambda}{n-1} \mathfrak{J}(b, w_1) + \dots + \lambda \mathfrak{J}(b, w_k) + \dots + \frac{1-\lambda}{n-1} \mathfrak{J}(b, w_n) \\
&\leq \frac{1-\lambda}{n-1} \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) + \dots + \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) + \dots + \frac{1-\lambda}{n-1} \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) \\
&= (1-\lambda) \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) + \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) \\
&= H_\lambda^{\mathfrak{J}}(b)
\end{aligned}$$

This completes the proof. □

Now, in order to prove Theorem 2, it suffices to show the following:

- (i) If b is probabilified by p_k^λ and b' is probabilified by p_j^λ , then

$$H_\lambda^{\mathfrak{J}}(b) = H_\lambda^{\mathfrak{J}}(b')$$

- (ii) If b is probabilified by p_k^λ and b' is not probabilified by p_j^λ for any $1 \leq j \leq n$, then

$$H_\lambda^{\mathfrak{J}}(b) < H_\lambda^{\mathfrak{J}}(b')$$

Proof of (i). First, suppose $j = k$. So b and b' are both probabilified by p_k^λ . Then, by Lemma 4(I),

$$H_\lambda^{\mathfrak{J}}(b) = \text{Exp}_{\mathfrak{J}}(b|p_k^\lambda) = \text{Exp}_{\mathfrak{J}}(b'|p_k^\lambda) = H_\lambda^{\mathfrak{J}}(b')$$

since, by Theorem 1, the expected inaccuracy of a belief function is minimal relative to a

probability function iff that probability function probabilifies it.

Second, suppose $j \neq k$. Note that p_k^λ is obtained from p_j^λ by a permutation of the worlds in $\mathcal{W}_{\mathcal{F}}$ and that there is a belief function that is probabilified by p_k^λ that is obtained from b' by the same permutation. Then note that $H_\lambda^\mathfrak{J}$ is invariant under such permutations.

Proof of (ii). Suppose $\mathfrak{J}(b, w_i) \leq \mathfrak{J}(b, w_j)$ for all $1 \leq j \leq n$. And suppose b^* is probabilified by p_i^λ . Then, by Theorem 1,

$$\text{Exp}_{\mathfrak{J}}(b^* | p_i^\lambda) < \text{Exp}_{\mathfrak{J}}(b' | p_i^\lambda)$$

By Lemma 4(II), we have $\text{Exp}_{\mathfrak{J}}(b' | p_i^\lambda) \leq H_\lambda^\mathfrak{J}(b')$. By (i) and Lemma 4(I), we have $\text{Exp}_{\mathfrak{J}}(b^* | p_i^\lambda) = H_\lambda^\mathfrak{J}(b^*) = H_\lambda^\mathfrak{J}(b)$. Putting these together, and appealing to Lemma 4(II), we have

$$H_\lambda^\mathfrak{J}(b) = H_\lambda^\mathfrak{J}(b^*) = \text{Exp}_{\mathfrak{J}}(b^* | p_i^\lambda) < \text{Exp}_{\mathfrak{J}}(b' | p_i^\lambda) = H_\lambda^\mathfrak{J}(b')$$

CASE II Next, let us assume that $\frac{1}{n} \geq \lambda$. We prove the following lemma:

Lemma 5 Suppose $\lambda \leq \frac{1}{n}$.

(I) If b is probabilified by p_k^λ , then

$$\text{Exp}_{\mathfrak{J}}(b | p_k^\lambda) = H_\lambda^\mathfrak{J}(b)$$

(II) If b is not probabilified by p_j^λ for any $1 \leq j \leq n$ then

$$\text{Exp}_{\mathfrak{J}}(b | p_k^\lambda) \leq H_\lambda^\mathfrak{J}(b)$$

Proof of Lemma 5(I). First, note that, if b is probabilified by p_k^λ , then $\mathfrak{J}(b, w_i) = \mathfrak{J}(b, w_j)$, for any $1 \leq i, j, \leq n$. So $\min_{w \in \mathcal{W}} \mathfrak{J}(b, w) = \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) = \mathfrak{J}(b, w_i)$, for any $1 \leq i \leq n$. Thus:

$$\begin{aligned} \text{Exp}_{\mathfrak{J}}(b | p_k^\lambda) &= \sum_{i=1}^n p_k^\lambda(w_i) \mathfrak{J}(b, w_i) \\ &= \frac{1}{n} \mathfrak{J}(b, w_1) + \dots + \frac{1}{n} \mathfrak{J}(b, w_n) \\ &= \mathfrak{J}(b, w_i) \text{ for any } 1 \leq i \leq n \\ &= \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) + (1 - \lambda) \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) \\ &= H_\lambda^\mathfrak{J}(b) \end{aligned}$$

Proof of Theorem 4(II).

$$\begin{aligned}
\text{Exp}_{\mathfrak{J}}(b|p_k^\lambda) &= \sum_{i=1}^n p_k^\lambda(w_i) \mathfrak{J}(b, w_i) \\
&= \frac{1}{n} \mathfrak{J}(b, w_1) + \dots + \frac{1}{n} \mathfrak{J}(b, w_n) \\
&\leq \frac{1}{n} \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) + \frac{n-1}{n} \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) \\
&\leq \lambda \min_{w \in \mathcal{W}} \mathfrak{J}(b, w) + (1-\lambda) \max_{w \in \mathcal{W}} \mathfrak{J}(b, w) \\
&= H_\lambda^{\mathfrak{J}}(b)
\end{aligned}$$

This completes the proof. □

Using the same arguments as above, we can infer:

(i') If b is probabilified by p_k^λ and b' is probabilified by p_j^λ , then

$$H_\lambda^{\mathfrak{J}}(b) = H_\lambda^{\mathfrak{J}}(b')$$

(ii') If b is probabilified by p_k^λ and b' is not probabilified by p_j^λ for any $1 \leq j \leq n$, then

$$H_\lambda^{\mathfrak{J}}(b) < H_\lambda^{\mathfrak{J}}(b')$$

And that completes our proof of Theorem 2. □

Theorem 3 is proved in exactly the same way, substituting a credal inaccuracy measure \mathfrak{J} generated by a strictly proper local inaccuracy measure for the inaccuracy measure over belief functions used above. Wherever we appeal to Theorem 1 in our proof of Theorem 2, we appeal to the strict propriety of \mathfrak{s} in our proof of Theorem 3.

References

- Buchak, L. (2014). *Risk and Rationality*. Oxford University Press.
- Clifford, W. K. (1877). The Ethics of Belief. *Contemporary Review*, 29:289–309.
- de Finetti, B. (1974). *Theory of Probability*, volume 1. Wiley, New York.
- Easwaran, K. (ms). Dr Truthlove, Or: How I Learned to Stop Worrying and Love Bayesian Probabilities.
- Easwaran, K. and Fitelson, B. (ta). Accuracy, Coherence, and Evidence. *Oxford Studies in Epistemology*, 5.

- Fitelson, B. (ms). *Coherence*. Oxford University Press.
- Greaves, H. and Wallace, D. (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, 115(459):607–632.
- Hempel, C. (1962). Deductive-Nomological vs. Statistical Explanation. In Feigl, H. and Maxwell, G., editors, *Minnesota Studies in the Philosophy of Science*, volume III, pages 98–169. University of Minnesota Press, Minneapolis.
- James, W. (1896). The Will to Believe. *The New World*, 5:327–347.
- Joyce, J. M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, 65(4):575–603.
- Joyce, J. M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In Huber, F. and Schmidt-Petri, C., editors, *Degrees of Belief*. Springer.
- Konek, J. (ta). Epistemic Conservativity and Imprecise Credence. *Philosophy and Phenomenological Research*.
- Levi, I. (1967). *Gambling with Truth*. Knopf, New York.
- Pettigrew, R. (2014). Accuracy, Risk, and the Principle of Indifference. *Philosophy and Phenomenological Research*.
- Pettigrew, R. (ms). *Accuracy and the Laws of Credence*. Oxford University Press, Oxford.
- Predd, J., Seiringer, R., Lieb, E. H., Osherson, D., Poor, V., and Kulkarni, S. (2009). Probabilistic Coherence and Proper Scoring Rules. *IEEE Transactions of Information Theory*, 55(10):4786–4792.