

Setup ●○	Coherence ○○	From Coherence to Consistency ○○	Stability Theory vs. MEEU ○○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- We assume our agent has a credence function $b(\cdot)$, which is *probabilistic*. Probabilism for $b(\cdot)$ can *itself* be motivated *via* EUT [26]. Here, I will take this to be *common ground*.
- We assume that our agent takes exactly one of three qualitative attitudes (B, D, S) toward each member of a finite agenda \mathcal{A} of (classical, possible worlds) propositions.
- We do *not* assume that these qualitative judgments can be *reduced* to $b(\cdot)$. But, we will use $b(\cdot)$ to derive a *rational coherence constraint* for qualitative judgment sets \mathbf{B} (on \mathcal{A}).
- This derivation requires both the agent's credence function $b(\cdot)$ and their *epistemic utility function* [12, 19, 23] $u(\cdot)$.
 - ☞ Following Easwaran [4, 6], we assume our agent cares *only* about whether their qualitative judgments are *accurate*.
- Specifically, our agent attaches some *positive* utility (r) with making an *accurate* judgment, and some *negative* utility ($-\omega$) with making an *inaccurate* judgment (where $\omega \geq r > 0$).

Branden Fitelson Belief & Credence: The view from naïve EUT 2

Setup ○●	Coherence ○○	From Coherence to Consistency ○○	Stability Theory vs. MEEU ○○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- Because suspensions are neither accurate nor inaccurate (*per se*), our agent will attach *zero* epistemic utility to suspensions $S(p)$, independently of the truth-value of p .
- Thus, we have the following piecewise definition of $u(\cdot, w)$.
$$u(B(p), w) \stackrel{\text{def}}{=} \begin{cases} -\omega & \text{if } p \text{ is false at } w \\ r & \text{if } p \text{ is true at } w \end{cases}$$

$$u(D(p), w) \stackrel{\text{def}}{=} \begin{cases} r & \text{if } p \text{ is false at } w \\ -\omega & \text{if } p \text{ is true at } w \end{cases}$$

$$u(S(p), w) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } p \text{ is false at } w \\ 0 & \text{if } p \text{ is true at } w \end{cases}$$
- With this *accuracy-centered* epistemic utility function in hand, we can derive a naïve EUT coherence requirement.

Branden Fitelson Belief & Credence: The view from naïve EUT 3

Setup ○○	Coherence ●○	From Coherence to Consistency ○○	Stability Theory vs. MEEU ○○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- To do so, we'll also need a *decision-theoretic principle*.
- Applications of EUT to grounding probabilism as a (synchronic) requirement for $b(\cdot)$ typically appeal to a *non-dominance* (in epistemic utility) principle [15, 27, 26].
- But, some authors apply an *expected epistemic utility maximization* (or *expected inaccuracy minimization*) principle to derive rational requirements [18, 11, 5, 25].

Coherence. An agent's belief set \mathbf{B} over an agenda \mathcal{A} should, from the point of view of their own credence function $b(\cdot)$, *maximize expected epistemic utility* (or *minimize expected inaccuracy*). That is, \mathbf{B} should maximize

$$EEU(\mathbf{B}, b) \stackrel{\text{def}}{=} \sum_{p \in \mathcal{A}} \sum_{w \in W} b(w) \cdot u(\mathbf{B}(p), w)$$

where $\mathbf{B}(p)$ is the agent's attitude toward p , and $W \stackrel{\text{def}}{=} \bigcup \mathcal{A}$.
- For now, we assume "*act-state independence*": $\mathbf{B}(p)$ and p are *b-independent* [10, 2, 1, 16]. We'll return to this issue.

Branden Fitelson Belief & Credence: The view from naïve EUT 4

Setup ○○	Coherence ○●	From Coherence to Consistency ○○	Stability Theory vs. MEEU ○○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- The consequences of **Coherence** are rather simple and intuitive. It is straightforward to prove the following result.

Theorem ([4]). An agent with credence function $b(\cdot)$ and qualitative judgment set \mathbf{B} over agenda \mathcal{A} satisfies **Coherence** *if and only if* for all $p \in \mathcal{A}$

$$B(p) \in \mathbf{B} \text{ iff } b(p) > \frac{\omega}{r+\omega},$$

$$D(p) \in \mathbf{B} \text{ iff } b(p) < \frac{r}{r+\omega},$$

$$S(p) \in \mathbf{B} \text{ iff } b(p) \in \left[\frac{r}{r+\omega}, \frac{\omega}{r+\omega} \right].$$
- ☞ In other words, **Coherence** entails *Lockean representability*, where the Lockean thresholds are determined by the way the agent (relatively) values accuracy vs. inaccuracy.
- This provides an elegant, EUT-based explanation of why Lockean representability is a rational requirement for agents with *both* credences *and* qualitative attitudes.
- Next, I will explain when **Coherence** entails *consistency*.

Branden Fitelson Belief & Credence: The view from naïve EUT 5

Setup ○○	Coherence ○○	From Coherence to Consistency ●○	Stability Theory vs. MEEU ○○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- Suppose our (naïve) agent has a belief set \mathbf{B}_n on a *minimal inconsistent* agenda of size n (e.g., $(n - 1)$ -ticket lottery).

Theorem. For all $n \geq 2$ and any probability function $\text{Pr}(\cdot)$, the $\text{Pr}(\cdot)$ -Lockean-representability of \mathbf{B}_n (with threshold t) entails deductive consistency of \mathbf{B}_n iff $t \geq \frac{n-1}{n}$.
- If we combine this with Easwaran’s **Coherence** theorem, we get the following result, regarding the conditions under which the **Coherence** of \mathbf{B}_n entails the consistency of \mathbf{B}_n .

Theorem. For all $n \geq 2$, an agent with an accuracy-centered utility function u , a credence function $b(\cdot)$, and a belief set \mathbf{B}_n , the **Coherence** of \mathbf{B}_n entails the consistency of \mathbf{B}_n iff

$$(\dagger) \quad w \geq (n - 1) \cdot r.$$

☞ Insisting that **Coherence** implies consistency (wrt \mathbf{B}_n) requires (naïve) agents to disvalue inaccuracy at least $(n - 1)$ times as much as they value accuracy.

Branden Fitelson Belief & Credence: The view from naïve EUT 6

Setup ○○	Coherence ○○	From Coherence to Consistency ●○	Stability Theory vs. MEEU ○○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- Of course, there will be *some* agents with epistemic utility functions u , which *do* satisfy (\dagger) . But, it is very odd (from a traditional Bayesian perspective) to *require* that such an agent’s epistemic utility function *must* satisfy (\dagger) .
- For example, in Lottery Paradox cases, we can make n as large as we like. And, the larger we make n , the stronger (and more implausible) the constraint (\dagger) becomes.

☞ $\therefore \mathbf{B}_n$ -consistency won’t be a *universal* MEEU-requirement. In other words, consistency *outstrips* the MEEU-theory of epistemic rationality. Leitgeb [17] defends an alternative.

- According to Leitgeb’s Stability Theory [17], a rational agent with credence function b (over a set of possible worlds W) believes a proposition p , viz., $B(p)$ iff p is entailed by some proposition B_W that is *p-stable*, where this is defined as:

p-stability. Given a probability model $\langle W, b(\cdot) \rangle$, a proposition $x \in W$ is *p-stable* iff $b(x \mid y) > 1/2$, for all $y \in W$ such that $x \& y \neq \perp$ and $b(y) > 0$.

Branden Fitelson Belief & Credence: The view from naïve EUT 7

Setup ○○	Coherence ○○	From Coherence to Consistency ○○	Stability Theory vs. MEEU ●○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- Basically, Leitgeb’s theory requires an agent to satisfy a *resilient* [30] version of the Lockean Thesis.
- As Leitgeb explains, his theory will require that rational agents have *consistent* (and *closed* \therefore *cogent*) belief sets (even belief sets \mathbf{B}_n in $(n - 1)$ -ticket Lottery Paradoxes).
- So, by our argument above, Stability Theory must *outstrip* MEEU-theory, which does *not* require consistency of \mathbf{B}_n (at least, this is not required for *every* MEEU-rational agent).
- Leitgeb’s theory has several (*prima facie*) odd consequences. I will focus on one problematic feature: the *violation* of *partition-invariance*. As we’ll see, Leitgeb’s theory is partition-sensitive in a particularly troubling way.
 - See Extras for some other (*prima facie*) odd consequences.
- A requirement on rational belief (or rational action) is *partition-invariant* (PI) iff its prescriptions do not depend on how the underlying space of possibilities is partitioned.

Branden Fitelson Belief & Credence: The view from naïve EUT 8

Setup ○○	Coherence ○○	From Coherence to Consistency ○○	Stability Theory vs. MEEU ●○○○○	Extras ○○	References
-------------	-----------------	-------------------------------------	------------------------------------	--------------	------------

- In the case of practical rationality (viz., rational action), many philosophers endorse (PI) as a *desideratum* [13, 7, 8, 20, 14].
- Savage’s theory [28] and standard causal decision theories [9, 30, 21, 31] are *partition-dependent*. This has led various authors [13, 7, 14] to endorse evidential decision theories.
- We defined **Coherence** “Savage-style,” and we assumed *act-state independence* (ASI) to ensure (PI). For our present examples (e.g., Lotteries) this is OK. *But*, see [10, 2, 16].¹
- Lin & Kelly [22] show: adding cogency to *any* non-trivial probabilistic acceptance rule for belief will entail *partition sensitivity* — even in Lottery cases (i.e., *even if* ASI obtains).

☞ And, Schurz [29] shows that all cogent *Lockean* theories (i.e., all *Stability* theories) must violate an even more plausible invariance constraint, which he calls *Independence*.

¹More generally, **Coherence** will satisfy (PI) if u satisfies following, for all partitions $\{X_i\}$ of W : $(\forall X_i) [u(\mathbf{B}(p), X_i) = \sum_{w \in W} b(w \mid X_i) \cdot u(\mathbf{B}(p), w)]$.

Branden Fitelson Belief & Credence: The view from naïve EUT 9

- This stronger kind of partition-sensitivity can be illustrated *via* one of Leitgeb’s own examples ([17, pp. 164–168], [3]).
- Let $\neg E \stackrel{\text{def}}{=} \text{Adams’s prediction of the secular acceleration of the moon}$, $T \stackrel{\text{def}}{=} \text{Newtonian theory (the part Adams used to predict } \neg E)$, $H \stackrel{\text{def}}{=} \text{the auxiliary hypotheses (e.g., negligibility of tidal friction) Adams used in his deduction of } \neg E \text{ from } T$.
- Suppose the following probability model $\langle W, b(\cdot) \rangle$ represents to the *epistemically rational degrees of belief* of a mid-19th century scientist (e.g., Adams), prior to learning E .

w_i	E	H	T	$b(w_i)$
w_1	F	T	T	27/50
w_2	F	F	T	171/500
w_3	F	T	F	29/500
w_4	F	F	F	1997/50000
w_5	T	F	T	9/500
w_6	T	F	F	1/500
w_7	T	T	F	3/50000
w_8	T	T	T	0

- As Dorling explains, Bayesian confirmation theory (BCT) implies the following two verdicts regarding this example (*pace* what Quine & Duhem might have said about the case).

(1) E weakly disconfirms T

$$b(T | E) = 0.897308 \approx 0.9 = b(T)$$

(2) E strongly disconfirms H

$$b(H | E) = 0.00299103 \ll 0.59806 = b(H)$$

- Leitgeb offers a qualitative analysis of Dorling’s example (in ST), in which the agent starts off (prior to learning E) believing H , T , and $\neg E$. Then, after learning E , the agent comes to *disbelieve* H , while believing both E and T .
- ☞ Crucial to Leitgeb’s analysis is his choice of $B_W = \{w_1\}$, which is the strongest p -stable set, relative to $\langle W, b(\cdot) \rangle$.
- While Leitgeb’s qualitative analysis of the Dorling example is interesting, it is threatened by the non-Independence of Stability Theory. Allow me to explain, *via* a similar example.

- Add a fourth atomic sentence “ C ” to the language, which expresses the proposition that a fair coin-toss came up “heads”. Here’s the resulting probability model $\langle W', b'(\cdot) \rangle$.

w'_i	E	H	T	C	$b'(w'_i)$
w'_1	F	T	T	T	27/100
w'_2	F	F	T	T	171/1000
w'_3	F	T	F	T	29/1000
w'_4	F	F	F	T	1997/100000
w'_5	T	F	T	T	9/1000
w'_6	T	F	F	T	1/1000
w'_7	T	T	F	T	3/100000
w'_8	T	T	T	T	0
w'_9	F	T	T	F	27/100
w'_{10}	F	F	T	F	171/1000
w'_{11}	F	T	F	F	29/1000
w'_{12}	F	F	F	F	1997/100000
w'_{13}	T	F	T	F	9/1000
w'_{14}	T	F	F	F	1/1000
w'_{15}	T	T	F	F	3/100000
w'_{16}	T	T	T	F	0

- The two BCT verdicts [(1), (2)] above *remain exactly the same* in this new model $\langle W', b'(\cdot) \rangle$, since C is b' -independent.
- ☞ But, the strongest p -stable proposition relative to this new model $\langle W', b'(\cdot) \rangle$ is now $B_{W'} = \{w'_1, w'_2, w'_9, w'_{10}\}$, which *does not entail* H . So, Leitgeb’s ST-analysis is *undermined*.

- Here’s a conjecture regarding one possible way of getting to something the resembles ST, using the machinery of EUT.

Conjecture. Let \mathcal{Y} be any set of W -propositions (with nonzero b -credence). If a belief set \mathbf{B} (on \mathcal{A}) maximizes

$$EEU_{\mathcal{Y}}(\mathbf{B}, b) \stackrel{\text{def}}{=} \sum_{p \in \mathcal{A}} \sum_{w \in W} b(w | \mathcal{Y}) \cdot u(\mathbf{B}(p), w)$$

for all $y \in \mathcal{Y}$, then \mathbf{B} is resiliently Lockean representable by $b(\cdot | \mathcal{Y})$, for each $y \in \mathcal{Y}$, with threshold $t = \frac{w}{r+w}$.

- If this conjecture is true, then “Stability Theory” emerges from “resilient expected epistemic utility maximization.”

- S_1 and S_2 share the same credence function $b_1 = b_2 = b$. But, they have very different belief states \mathbf{B}_1 and \mathbf{B}_2 [24]. The following table depicts b , \mathbf{B}_1 and \mathbf{B}_2 (on the contingent p 's).

w 's	p	b	\mathbf{B}_1 (MEEU $_{1/2}$)	\mathbf{B}_2 (ST $_{1/2}$)
{ w_1 }	$\neg X \wedge \neg Y$	0.5	S	S
{ w_2 }	$X \wedge \neg Y$	0.25	D	S
{ w_3 }	$X \wedge Y$	0.125	D	S
{ w_4 }	$\neg X \wedge Y$	0.125	D	S
{ w_1, w_2 }	$\neg Y$	0.75	B	S
{ w_1, w_3 }	$X \equiv Y$	0.625	B	S
{ w_1, w_4 }	$\neg X$	0.625	B	S
{ w_2, w_3 }	X	0.375	D	S
{ w_2, w_4 }	$X \neq Y$	0.375	D	S
{ w_1, w_4 }	Y	0.25	D	S
{ w_1, w_2, w_3 }	$X \vee \neg Y$	0.875	B	S
{ w_1, w_2, w_4 }	$\neg X \vee \neg Y$	0.875	B	S
{ w_1, w_3, w_4 }	$\neg X \vee Y$	0.75	B	S
{ w_2, w_3, w_4 }	$X \vee Y$	0.5	S	S

- There are (arbitrarily) small perturbations b' of b , which (a) do not alter the $1/2$ -credence p 's, (b) lower the credence of $\neg X \vee \neg Y$, but (c) make it rational for S_2 to believe $\neg X \vee \neg Y$.

w 's	p	b	b'	$\mathbf{B}_1 = \mathbf{B}'_1$	\mathbf{B}_2	\mathbf{B}'_2
{ w_1 }	$\neg X \wedge \neg Y$	0.5	0.5	S	S	S
{ w_2 }	$X \wedge \neg Y$	0.25	0.2366	D	S	S
{ w_3 }	$X \wedge Y$	0.125	0.1295	D	S	D
{ w_4 }	$\neg X \wedge Y$	0.125	0.1339	D	S	S
{ w_1, w_2 }	$\neg Y$	0.75	0.7366	B	S	S
{ w_1, w_3 }	$X \equiv Y$	0.625	0.6295	B	S	S
{ w_1, w_4 }	$\neg X$	0.625	0.6339	B	S	S
{ w_2, w_3 }	X	0.375	0.3660	D	S	S
{ w_2, w_4 }	$X \neq Y$	0.375	0.3705	D	S	S
{ w_1, w_4 }	Y	0.25	0.2634	D	S	S
{ w_1, w_2, w_3 }	$X \vee \neg Y$	0.875	0.8661	B	S	S
{ w_1, w_2, w_4 }	$\neg X \vee \neg Y$	0.875	0.8705	B	S	B
{ w_1, w_3, w_4 }	$\neg X \vee Y$	0.75	0.7634	B	S	S
{ w_2, w_3, w_4 }	$X \vee Y$	0.5	0.5	S	S	S

- [1] M. Caie, *Rational Probabilistic Incoherence*, *Philosophical Review*, 2013.
- [2] J. Carr, *Epistemic Utility Theory and the Aim of Belief*, 2014.
- [3] J. Dorling, *Bayesian personalism, the methodology of scientific research programmes, and Duhem's problem*, *Studies in History and Philosophy of Science Part A*, 1979.
- [4] K. Easwaran, *Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probability*, manuscript, September 2014.
- [5] ———, *Expected Accuracy Supports Conditionalization—and Conglomerability and Reflection*, *Philosophy of Science*, 2013.
- [6] K. Easwaran and B. Fitelson, *Accuracy, Coherence, and Evidence*, to appear in *Oxford Studies in Epistemology V*, T. Szabo-Gendler and J. Hawthorne (eds.).
- [7] E. Eells, *Rational Decision and Causality*, CUP, 1982.
- [8] ———, *Levi's "The Wrong Box"*, *Journal of Philosophy*, 1985.
- [9] A. Gibbard & W. Harper. *Counterfactuals and two kinds of expected utility*, 1981.
- [10] H. Greaves, *Epistemic Decision Theory*, *Mind*, 2013.
- [11] H. Greaves and D. Wallace, *Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility*, *Mind*, 2006.
- [12] C. Hempel, *Deductive-Nomological vs. Statistical Explanation*, 1962.
- [13] R. Jeffrey, *The Logic of Decision*, Chicago, 1990.
- [14] J. Joyce, *The Foundations of Causal Decision Theory*, CUP, 1999.

- [15] ———, *Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief*, in F. Huber & C. Schmidt-Petri (eds.), *Degrees of Belief*, 2009.
- [16] J. Konek and B. Levinstein, *The Foundations of Epistemic Decision Theory*, 2014.
- [17] H. Leitgeb, *The Stability Theory of Belief*, *Philosophical Review*, 2014.
- [18] H. Leitgeb and R. Pettigrew, *An Objective Justification of Bayesianism I & II*, *Philosophy of Science*, 2010.
- [19] I. Levi, *Gambling with Truth*, 1967.
- [20] ———, *The Wrong Box*, *Journal of Philosophy*, 1985.
- [21] D. Lewis, *Causal decision theory*, *Australasian Journal of Philosophy*, 1981.
- [22] H. Lin and K. Kelly, *A Geo-logical Solution to the Lottery Paradox*, 2012.
- [23] P. Maher, *Betting on Theories*, 1993.
- [24] D. Makinson, *Remarks on the Stability Theory of Belief*, 2014.
- [25] G. Oddie, *Conditionalization, cogency, and cognitive value*, *British Journal for the Philosophy of Science*, 1997.
- [26] R. Pettigrew, *Epistemic Utility Arguments for Probabilism*, *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), URL = <http://plato.stanford.edu/archives/win2011/entries/epistemic-utility/>.
- [27] J. Predd, R. Seringer, E. Loeb, D. Osherson, H.V. Poor and S. Kulkarni, *Probabilistic coherence and proper scoring rules*, *IEEE Transactions*, 2009.
- [28] L. Savage, *The Foundations of Statistics*, Dover, 1972.
- [29] G. Schurz, *Impossibility Results for Stable Belief*, 2014.
- [30] B. Skyrms, *Causal decision theory*, *Journal of Philosophy*, 1982.
- [31] J.H. Sobel, *Taking Chances: Essays on Rational Choice*, 1994.