

Two Approaches to Belief Revision

Ted Shear¹ Jonathan Weisberg² Branden Fitelson³

¹University of California, Davis

ebshear@ucdavis.edu

²University of Toronto

jonathan.weisberg@utoronto.ca

³Northeastern University

branden@fitelson.org

Formal Epistemology Workshop

June 20, 2016

Today, we will explore a new approach to belief revision—EUT revision—motivated in terms of *epistemic utility theory*

To aid in this exploration, we will contrast it with the orthodox approach belief revision provided by the AGM theory.

While AGM revision is driven by purely logical considerations, EUT revision involves a (normative) Lockean thesis and results in a broadly Bayesian approach to the revision of full beliefs.

In addition to qualitative belief sets, \mathbf{B} , our agents possess a numerical credence functions, $b(\cdot)$.

On the belief side, our agents entertain (classical, possible worlds) propositions on some finite *agenda* \mathcal{A} .

- (1) \mathbf{B} is the set of propositions in \mathcal{A} believed by our agent.
 - *Note:* when $p \in \mathbf{B}$, we write $\mathbf{B}(p)$.
- (2) Given a *prior* belief set \mathbf{B} , the *posterior* belief set \mathbf{B}' is generated by revising the prior by E — *i.e.*, $\mathbf{B}' = \mathbf{B} \star E$.

On the credence side:

- (3) $b(\cdot)$ is a classical (Kolmogorov) probability function.
- (4) Given a *prior* $b(\cdot)$, the *posterior* $b'(\cdot)$ is generated via *conditionalization* by E — *i.e.*, $b'(\cdot) = b(\cdot | E)$.¹

¹Our results generalize to any “minimum distance” [4] Bayesian update satisfying (i) $b'(E) > b(E)$, (ii) $b'(E) > t$, and (iii) $b(E \supset X) \geq b'(X)$ will suffice.

Plan for the talk:

- ① explain EUT basics,
- ② explain AGM basics,
- ③ assess EUT under the AGM axioms, and
- ④ present some further interesting results.

[Stay awake to hear the shocking (!) mystery result.]

Ultimately, the EUT approach will generate a Lockean thesis:

$$\mathbf{B}(p) \quad \text{iff} \quad b(p) > t.$$

Of course, accounts based solely on the Lockean thesis permit belief sets that are not *cogent*.

Cogency. An agent's belief set \mathbf{B} should (at any given time) be deductively consistent and closed under logic.

We will avoid retrodding worn paths by focusing on *cogent* agents.

EUT's core principle [12, 16]:

- At any given time, an agent should have beliefs that *maximize expected epistemic utility* relative to her credence function.

We take a *veritistic* approach to epistemic utility according to which the only value-relevant feature of belief is *accuracy*.²

$$u(\mathbf{B}(p), w) := \begin{cases} r & \text{if } p \text{ is true at } w \\ -w & \text{if } p \text{ is false at } w \end{cases}$$

The only constraint imposed on r and w is:

$$(\dagger) \quad w > \left(\frac{1 + \sqrt{5}}{2} \right) \cdot r > 0.$$

This peculiar looking constraint and the surprising presence of the Golden ratio ($\Phi = \frac{1 + \sqrt{5}}{2}$) in it will become clearer as we progress.

²The epistemic utility of suspending belief on p is 0 (*i.e.* when $p, \neg p \notin \mathbf{B}$).

The *expected epistemic utility (EEU)* of a *belief* $\mathbf{B}(p)$, from the point of view of a credence function $b(\cdot)$, is calculated as normal.

$$EEU(\mathbf{B}(p), b) := \sum_{w \in W} b(w) \cdot u(\mathbf{B}(p), w)$$

The *overall EEU* of an agent's *belief set* \mathbf{B} is simply the sum of the EEUs of their beliefs.

$$EEU(\mathbf{B}, b) := \sum_{p \in \mathbf{B}} EEU(\mathbf{B}(p), b)$$

This generates EUT's Lockean thesis:

Theorem (Dorst [5], Easwaran [7]). A belief set, \mathbf{B} , *maximizes EEU relative to* b just in case, for every $p \in \mathcal{A}$:

$$\mathbf{B}(p) \quad \text{iff} \quad b(p) > \frac{w}{r + w}.$$

Recall EUT's core principle:

- *At any given time*, an agent should have beliefs that *maximize* her *expected epistemic utility* relative to her credence function at that time.

This immediately generates a diachronic norm on belief:

- When an agent revises her beliefs, her posterior belief set \mathbf{B}' should maximize *EEU* relative to her posterior credence function, b' .

Since $b'(X) = b(X | E)$, this entails that $\mathbf{B}' = \mathbf{B} * E$.

$$\mathbf{B} * E := \left\{ p \mid b(p | E) > \frac{w}{r+w} \right\}$$

To get a feel for how EUT revision works, observe that EUT does not satisfy the following principle:

(P2) If an agent learns something that she *already* believes, then her belief set should *remain unchanged*.

Or: $\mathbf{B} \star X = \mathbf{B}$, for all $X \in \mathbf{B}$.

Informally, the reason \star may violate (P2) is that conditionalizing by a proposition — even one above some threshold — can drop the credence assigned to some other proposition to below the threshold.

Proposition. If $b(p) > \frac{w}{r+w}$ and $b(q) > \frac{w}{r+w}$, then $b(p | q) > \frac{w-r}{w}$.

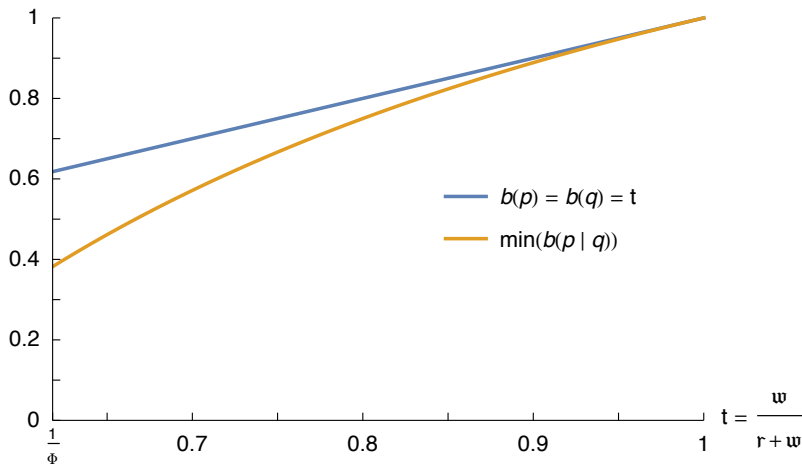


Figure: Bound on the “degree” that (P2) can fail

Corollary. If $B(p)$ and $B(q)$, then $\neg p \notin \mathbf{B} * q$.

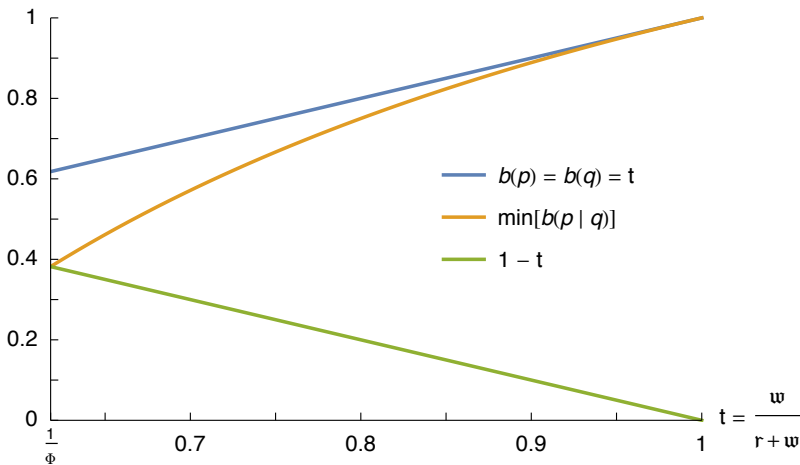


Figure: (P2) failures never results in disbelief of a prior belief

Corollary. * satisfies (P2), if there is no $p : b(p) \in [t, 1 - t + t^2]$.

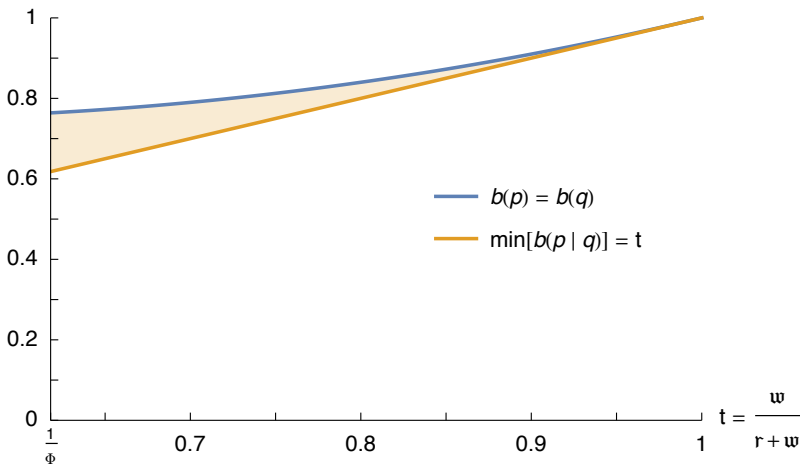


Figure: (P2)'s "Lockean danger-zone"

AGM theory is the orthodox account of belief revision and is very modest in the restrictions that it places on revisions.

The principle unifying AGM revision is *the principle of conservativity* (also called *informational economy* or *minimal mutilation*).

Conservativity. When an agent with a prior belief set \mathbf{B} learns E , she should revise to a posterior belief set that:

- ① *includes* E ,
- ② *is cogent*, and
- ③ *is a minimal change*³ to \mathbf{B} .

AGM revision has a variety of nice mathematical properties that have made for fruitful investigation.

Specifically, it has a simple and straightforward axiomatization.

³Minimal change can be thought of in terms of closeness between belief sets using Hamming distance, or any of a variety of other distance measures [13, 3, 6].

Basic Gärdenfors postulates:

- (*1) $\mathbf{B} * E = \text{Cn}(\mathbf{B} * E)$ **Closure**
- (*2) $E \in \mathbf{B} * E$ **Success**
- (*3) $\mathbf{B} * E \subseteq \text{Cn}(\mathbf{B} \cup \{E\})$ **Inclusion**
- (*4) If E is consistent with \mathbf{B} , then $\mathbf{B} * E \supseteq \text{Cn}(\mathbf{B} \cup \{E\})$ **Vacuity**
- (*5) If E is not a contradiction, then $\mathbf{B} * E$ is consistent **Consistency**
- (*6) If $X \Leftrightarrow Y$, then $\mathbf{B} * X = \mathbf{B} * Y$ **Extensionality**

Supplementary postulates:

- (*7) $\mathbf{B} * (E \wedge E') \subseteq \text{Cn}((\mathbf{B} * E) \cup \{E'\})$ **Superexpansion**
- (*8) If E' is consistent with $\text{Cn}(\mathbf{B} * E)$, then
 $\mathbf{B} * (E \wedge E') \supseteq \text{Cn}((\mathbf{B} * E) \cup \{E'\})$ **Subexpansion**

The following axiom will be useful in a few ways:

$$(*\top) \mathbf{B} * \top = \mathbf{B}$$

Idempotence

It helps explain why we view closure and consistency as standing synchronic constraints of AGM.

- Consider the consistent, but not closed, belief set $\mathbf{B} = \{p\}$.
By **Closure**, $\mathbf{B} * \top \neq \mathbf{B}$.
- Consider the closed, but *inconsistent*, belief set $p, \neg p \in \mathbf{B}$.
By **Consistency**, $\mathbf{B} * \top \neq \mathbf{B}$.

When **Idempotence** is accepted along with the other axioms, **Superexpansion** and **Subexpansion** respectively entail **Inclusion** and **Vacuity** (see Extras slide 1 for proof).

(*3) $\mathbf{B} \star E \subseteq \mathbf{Cn}(\mathbf{B} \cup \{E\})$

Inclusion

Proposition. \star satisfies **Inclusion**.

Proof: Let $X \in \mathbf{B} \star E$. So, $b(X | E) > \frac{w}{r+w}$. This implies that $b(E \supset X) > \frac{w}{r+w}$ since $\Pr(E \supset X) \geq \Pr(X | E)$. Thus, $E \supset X \in \mathbf{B}$ and by *modus ponens* we arrive at $X \in \mathbf{Cn}(\mathbf{B} \cup \{E\})$. \square

(*7) $\mathbf{B} * (E \wedge E') \subseteq \mathbf{Cn}((\mathbf{B} * E) \cup \{E'\})$

Superexpansion

Proposition. $*$ satisfies **Superexpansion**.

Proof: Let $X \in \mathbf{B} * (E \wedge E')$. So, $b(X | E \wedge E') > \frac{w}{r+w}$. Again, since $\Pr((E \wedge E') \supset X) \geq \Pr(X | E \wedge E')$, we know $b((E \wedge E') \supset X) > \frac{w}{r+w}$. So, $(E \wedge E') \supset X \in \mathbf{B}$. Additionally, we have $E \in \mathbf{B} * E$ by **Success**. Hence, by *modus ponens* $X \in \mathbf{Cn}((\mathbf{B} * E) \cup \{E'\})$. \square

(*4) If E is consistent with \mathbf{B} , then $\mathbf{B} \star E \supseteq \text{Cn}(\mathbf{B} \cup \{E\})$

Vacuity

Proposition. \star does *not* satisfy **Vacuity**.

Proof: Consider a simple random sampling task from the *urn model* displayed to the right.

E = ‘The object sampled will be red’

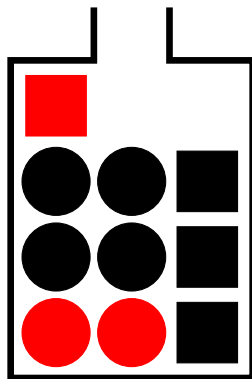
X = ‘The object sampled will be a circle’

Let $w = 0.17$ and $r = 0.03$, so that $\frac{w}{r+w} = 0.85$.

$E \supset X$ is only proposition with a probability higher than 0.85, so the EUT agent’s prior is:

$$\mathbf{B} = \{E \supset X\}.$$

But, what happens if she learns E ?



(*4) If E is consistent with \mathbf{B} , then $\mathbf{B} \star E \supseteq \text{Cn}(\mathbf{B} \cup \{E\})$

Vacuity

Proposition. \star does *not* satisfy **Vacuity**.

Proof: (continued)

Now, having learned E consistent with \mathbf{B} ,

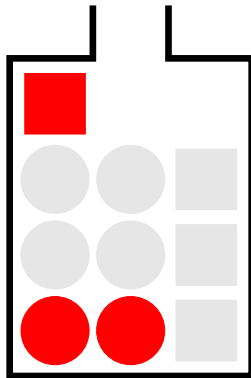
$$b(E \supset X | E) = 2/3 < 0.85.$$

Thus, $E \supset X \notin \mathbf{B} \star E$, but $E \supset X \in \mathbf{B}$. □

What will the EUT agent believe in this case?

$$\mathbf{B} \star E = \{E, E \vee X, E \vee \neg X\}.$$

So, the EUT agent is responsive to the non-definitive counter-evidence against her belief that $E \supset X$, while the AGM agent will accept the additional beliefs by *modus ponens*.



We will finish by looking at two interesting features of EUT's failures of **Vacuity** that both suggest that EUT should be viewed as more *epistemically risk-averse* than AGM.⁴

Theorem. $\mathbf{B} * E$ violates **Vacuity** iff $\mathbf{B} * E \subset \mathbf{B} * E$, for any E consistent with \mathbf{B} .

Proof: See Extras Slide 5

So, every time that EUT and AGM diverge, AGM is more “epistemically demanding” on the agent's beliefs; so AGM may be seen to be more *epistemically risk-seeking*.

⁴Pettigrew [17] has independently argued (*via* the use of an epistemic *Hurwicz Criterion*) that **Cogency** implies its own variety of *risk-seeking*.

Our final theorem, identifies an additional feature of EUT's violations of **Vacuity**.

For now, forget the constraint † on EUT's threshold and let $t \in [1/2, 1)$.

Amazingly, the lowest possible Lockean threshold permitting counterexamples to **Vacuity** for cogent agents is $t = \Phi^{-1} \approx 0.618$.⁵

Theorem. An EUT revision on any *cogent* prior may violate **Vacuity** only if $t \in [\Phi^{-1}, 1)$.

An immediate corollary of this extends the well-known convergence between AGM and *extremal* Lockean agents [10, 11].

Corollary. If $t \notin [\Phi^{-1}, 1)$, then EUT revision *is* an AGM revision for cogent agents.

⁵Moreover, counterexamples near Φ^{-1} are only possible for algebras with 3 strongest worlds; the bound is raised to $\frac{1}{\sqrt{2}}$ as soon as we move to algebras with 4 strongest worlds.

- The holy grail for this project is to find a purely qualitative characterization/axiomatization of EUT revision.
- Jan van Eijck & Bryan Renne [8] recently provided a modal logic for belief given a Lockean threshold of $1/2$, which we are hoping to use in characterizing EUT belief revision for a threshold of $1/2$.
- Our final theorem suggests that this may also be manageable for a threshold of Φ^{-1} .
- Since both EUT [4] and AGM [13] may be described terms of “minimal distance” revision, this yields a general “geodesic update” framework in which we plan to investigate *contraction* and other kinds of revision.

- [1] C. Alchourron, P. Gärdenfors, and D. Makinson. *On the logic of theory change: Partial meet contraction and revision functions*, 1985.
- [2] D. Christensen, *Putting Logic in its Place*, 2007.
- [3] M. Deza and E. Deza, *Encyclopedia of Distances*, 2009.
- [4] P. Diaconis and S. Zabell, *Updating Subjective Probability*, 1982.
- [5] K. Dorst, *Lockeans Maximize Expected Accuracy*, 2015.
- [6] C. Duddy and A. Piggins, *A measure of distance between judgment sets*, 2012.
- [7] K. Easwaran, *Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probability*, 2015.
- [8] J. van Eijck and B. Renne, *Belief as Willingness to Bet*, 2014.
- [9] R. Foley, *Working Without a Net*, 1992.
- [10] W. Harper, *Rational belief change, popper functions and counterfactuals*, 1975.
- [11] J. Hawthorne, *A primer on rational consequence relations, popper functions, and their ranked structure*, 2013.
- [12] C. Hempel, *Deductive-nomological vs. statistical explanation*, 1962.
- [13] K. Georgatos, *Geodesic Revision*, 2008.
- [14] H. Leitgeb, *The review paradox: On the diachronic costs of not closing rational belief under conjunction*, 2013.
- [15] Hannes Leitgeb. *Stability Theory of Belief*. Oxford University Press, 2016.
- [16] I. Levi, *Gambling with Truth*, 1967.
- [17] R. Pettigrew, *Accuracy and Risk: a Jamesian investigation in formal epistemology*, 2015.

Thanks!

Ted Shear

ebshear@ucdavis.edu

Jonathan Weisberg

jonathan.weisberg@utoronto.ca

Branden Fitelson

branden@fitelson.org

Draft & slides available at:

tedshear.com/research

Given the other AGM axioms, **Superexpansion** and **Subexpansion** imply **Inclusion** and **Vacuity**, respectively, assuming only the following weak additional postulate.

- | | |
|---|------------------------------|
| 1. $Y \in \mathbf{B} * X$ | Assumption |
| 2. $\mathbf{B} * X = \mathbf{B} * (\top \wedge X)$ | (1), Extensionality |
| 3. $Y \in \mathbf{B} * (\top \wedge X)$ | (1), (2) |
| 4. $Y \in \text{Cn}((\mathbf{B} * \top) \cup \{X\})$ | (3), Superexpansion |
| 5. $\text{Cn}((\mathbf{B} * \top) \cup \{X\}) = \text{Cn}(\mathbf{B} \cup \{X\})$ | Idempotence |
| 6. $Y \in \text{Cn}(\mathbf{B} \cup \{X\})$ | (4), (5) □ |
| | |
| 1. X is consistent with \mathbf{B} | Assumption |
| 2. $Y \in \text{Cn}(\mathbf{B} \cup \{X\})$ | Assumption |
| 3. $Y \in \text{Cn}((\mathbf{B} * \top) \cup \{X\})$ | (2), Idempotence |
| 4. $Y \in \mathbf{B} * (\top \wedge X)$ | (3), Subexpansion |
| 5. $Y \in \mathbf{B} * X$ | (4), Extensionality □ |

Alternative Axiomatization of AGM, using **Idempotence**.

- (*1) $\mathbf{B} * E = \text{Cn}(\mathbf{B} * E)$ **Closure**
- (*2) $E \in \mathbf{B} * E$ **Success**
- (*5) If E is not a contradiction, then $\mathbf{B} * E$ is consistent **Consistency**
- (*6) If $E \Leftrightarrow E'$, then $\mathbf{B} * E = \mathbf{B} * E'$ **Extensionality**
- (*7) $\mathbf{B} * (E \wedge E') \subseteq \text{Cn}((\mathbf{B} * E) \cup \{E'\})$ **Superexpansion**
- (*8) If E' is consistent with $\text{Cn}(\mathbf{B} * E)$, then
 $\mathbf{B} * (E \wedge E') \supseteq \text{Cn}((\mathbf{B} * E) \cup \{E'\})$ **Subexpansion**
- (*9) $\mathbf{B} * \top = \mathbf{B}$ **Idempotence**

- | | |
|---|--------------------------------|
| 1. \mathbf{B} is consistent. | Assumption |
| 2. \mathbf{B} is closed, <i>i.e.</i> , $\mathbf{B} = \text{Cn}(\mathbf{B})$. | Assumption |
| 3. $X \in \mathbf{B}$. | Assumption |
| 4. X is consistent with \mathbf{B} . | (1), (3), Logic |
| 5. $\mathbf{B} * X = \text{Cn}(\mathbf{B} \cup \{X\})$. | (4), Vacuity, Inclusion |
| 6. $\mathbf{B} * X = \text{Cn}(\mathbf{B})$. | (5), (3), Logic |
| 7. $\mathbf{B} * X = \text{Cn}(\mathbf{B} * X)$ | Closure |
| 8. $\text{Cn}(\mathbf{B} * X) = \text{Cn}(\mathbf{B})$ | (6), (7), Logic |
| 9. $\mathbf{B} * X = \mathbf{B}$ | (7), (8), (2), Logic \square |

Figure: Derivation of (P2) from **Closure**, **Inclusion**, and **Vacuity**

p	$b(p)$	$b(p E)$	$p \in \mathbf{B}?$	$p \in \mathbf{B} * E?$	$p \in \mathbf{B} \ast E?$	$p \in \text{Cn}(\mathbf{B} \cup \{E\})?$
$E \wedge X$	2/10	2/3	No	No	Yes	Yes
$E \wedge \neg X$	1/10	1/3	No	No	No	No
$\neg E \wedge X$	4/10	0	No	No	No	No
$\neg E \wedge \neg X$	3/10	0	No	No	No	No
E	3/10	1	No	Yes	Yes	Yes
X	6/10	2/3	No	No	Yes	Yes
$E \equiv X$	5/10	2/3	No	No	Yes	Yes
$E \equiv \neg X$	5/10	1/3	No	No	No	No
$\neg E$	7/10	0	No	No	No	No
$\neg X$	4/10	1/3	No	No	No	No
$E \vee X$	7/10	1	No	Yes	Yes	Yes
$E \vee \neg X$	6/10	1	No	Yes	Yes	Yes
$\neg E \vee X$	9/10	2/3	Yes	No	No	Yes
$\neg E \vee \neg X$	8/10	1/3	No	No	No	No

Table: Full probability distribution for EUT **Vacuity** counterexample

Theorem. $\mathbf{B} \ast E$ violates **Vacuity** iff $\mathbf{B} \ast E \subset \mathbf{B} \ast E$, for any E consistent with \mathbf{B} .

Proof: (\Rightarrow) Suppose EUT violates **Vacuity** (wrt \mathbf{B} and E). Then, (a) E is consistent with \mathbf{B} ; and, (b) $\mathbf{B} \ast E \not\subseteq \text{Cn}(\mathbf{B} \cup \{E\})$. By (b), there exists an X such that $X \in \text{Cn}(\mathbf{B} \cup \{E\})$ but $X \notin \mathbf{B} \ast E$. It follows from (a), **Vacuity** and **Inclusion** that $\text{Cn}(\mathbf{B} \cup \{E\}) = \mathbf{B} \ast E$. Therefore, $X \in \mathbf{B} \ast E$ and $X \notin \mathbf{B} \ast E$. And, by **Inclusion**, $\mathbf{B} \ast E \subseteq \text{Cn}(\mathbf{B} \cup \{E\}) = \mathbf{B} \ast E$. \square

(\Leftarrow) Suppose E is consistent with \mathbf{B} and $\mathbf{B} \ast E \subset \mathbf{B} \ast E$. Then, there exists an X such that $X \in \mathbf{B} \ast E$ but $X \notin \mathbf{B} \ast E$. Because E is consistent with \mathbf{B} , **Vacuity** and **Inclusion** imply that $\mathbf{B} \ast E = \text{Cn}(\mathbf{B} \cup \{E\})$. Therefore, $X \in \text{Cn}(\mathbf{B} \cup \{E\})$; but, $X \notin \mathbf{B} \ast E$. \square

Leitgeb [15] provides a number of remarkable representation theorems for his so-called *Humean thesis* — some may seem at odds with the present results.

The Humean thesis is as follows (where $r > 1/2$):

(HT r) $\mathbf{B}(X)$ iff $b(X | Y) > r$ for all Y such that $\neg Y \notin \mathbf{B}$ and $b(Y) > 0$.

Theorem ([15]). An agent's belief set, \mathbf{B} , and credence function, b , jointly satisfy (HT r) just in case:

- (a) \mathbf{B} is cogent,
- (b) b is a probabilistic credence function, and
- (c) $\mathbf{B}(p)$ iff $b(p) > r'$,

Note: The Humean threshold r (in HT r) is *not* the same as the Lockean threshold r' in (c).

Leitgeb offers the following belief revision postulates — which together yield AGM revision — on the way towards a diachronic representation theorem:

- (\circ BP1) For all $b(Y) > 0$, if $Z \in \mathbf{B} \circ Y$, then $b(Z | Y) > r$
- (\circ BP2) $b(Y) = 0$ iff $\mathbf{B} \circ Y$ is inconsistent
 - (\circ 1) $X \in \mathbf{B} \circ X$
 - (\circ 2) If $Y \in \mathbf{B} \circ X$ and $Y \vdash Z$, then $Z \in \mathbf{B} \circ X$
 - (\circ 3) If $Y \in \mathbf{B} \circ X$ and $Z \in \mathbf{B} \circ X$, then $Y \wedge Z \in \mathbf{B} \circ X$
 - (\circ 4) For any $\mathcal{Y} = \{Y | Y \in \mathbf{B} \circ X\}$, $\bigwedge \mathcal{Y} \in \mathbf{B} \circ X$
 - (\circ 5) $\perp \notin \mathbf{B} \circ \top$
 - (\circ 6) If Y is consistent with $\mathbf{B} \circ X$, then $\mathbf{B} \circ (X \wedge Y) = \text{Cn}(\mathbf{B} \circ X \cup \{Y\})$

One final definition:

P -stability r . X is P -stable r iff $b(X | Y) > r$ for any Y such that Y is consistent with X and $b(Y) > 0$.

This leads us to Leitgeb's relevant representation theorem:

Theorem ([15]). A revision operator \circ satisfies \circ BP1, \circ BP2, and $(\circ 1) - (\circ 6)$ iff *there is some class, \mathcal{X} , of non-empty P -stable^r propositions whose strongest member is X such that:*

- \mathcal{X} contains the least set of probability 1 in the algebra,
- all other members of \mathcal{X} have probability less than 1,
- for all Y , if $b(Y) > 0$ and Y is consistent with X , then $\mathbf{B} \circ Y = \mathbf{Cn}(Y \cup X)$, and
- for all Y , if $b(Y) = 0$, then $\mathbf{B} \circ Y$ is inconsistent.

Note: The theorem merely depends on the existence of *some* P -stable^r set — there is no guarantee that Leitgeb's preferred revision involves the *only* P -stable^r set.