Is Bayesian Coherentism Impossible?

Jonah N. Schupbach

April 23, 2007

Abstract

Bovens and Hartmann present an "impossibility result" against Bayesian Coherentism. This result putatively shows that coherence is separable if and only if it cannot be given a probabilistic, complete and transitive ordering relation. Bovens and Hartmann intend their result to apply to any such ordering, and thus to any proposed order-inducing probabilistic measure of coherence. Underlying their notion of separability - and thus underlying their impossibility result - is Bovens and Hartmann's introduction and support of a set of specific ceteris paribus conditions. In this paper, I argue that these ceteris paribus conditions are not clearly appropriate. Certain proposed coherence measures not only motivate different such conditions but they also call for the rejection of at least one of Bovens and Hartmann's conditions. I show that there exist sets of ceteris paribus conditions which, at least prima facie, have the same intuitive advantages as Bovens and Hartmann's conditions but which also allow one to sidestep the impossibility result altogether. This shifts the debate from the merits of the impossibility result itself to the underlying choice of ceteris paribus conditions.

B&H on Bayesian Coherentism

In several recent publications [(2003), (2005), (2006)], Luc Bovens and Stephan Hartmann (B&H) present their "impossibility result" for Bayesian Coherentism. They understand Bayesian Coherentism essentially to be the conjunction of the following three fundamental tenets [(2006, 78-9), (2003, 11-12, 25)]:

Separability: (BC_1): For all information sets S, $S' \in \boldsymbol{S}$, if S is no less coherent than S′ ($S \succcurlyeq S'$), then our degree of confidence that the content of S is true is no less than our degree of confidence that the content of S′ is true, ceteris paribus.

Probabilism: ($BC_{2[i]}$): The binary relation of "...being no less coherent than..." [i.e., $\succcurlyeq$] over $\boldsymbol{S}$ is fully determined by the probabilistic features of the information sets contained in $\boldsymbol{S}$.

Ordering: ($BC_{2[ii]}$): The binary relation of "...being no less coherent than..." [i.e., $\succcurlyeq$] is an ordering; i.e., the relation $\succcurlyeq$ is transitive and complete.[1]

[1] Following B&H's lead, I will refer to the conjunction of ($BC_{2[i]}$) and ($BC_{2[ii]}$) as (BC_2).

Importantly, B&H propose specific ceteris paribus conditions to enforce in the testing for separability. Framing the issue of coherence within a testimonial context that assumes independence of witnesses as a general prerequisite,[2] they suggest that there are three distinct "factors that affect our degree of confidence:"

1. Expectedness of the information: $P(R_1, R_2, ..., R_n)$ where R_i represents the i'th "information item" in the information set.[3]

2. Reliability of the information sources: $r := 1 - \bar{r} = 1 - q/p = 1 - P(REPR_i|\neg R_i)/P(REPR_i|R_i)$ where $REPR_i$ represents the positive value for the report variable on information item R_i; $REPR_i$ should be read, "a report is received to the effect that R_i."

3. Coherence of the information.

Given this setup, B&H propose that in order to meet separability's ceteris paribus conditions, one must hold factors (1) and (2) constant between sets. That is, in order to detect the effects of a difference in coherence between sets, one needs to hold all else equal, and this "all else" - per B&H - is made up of the expectedness of the information and the reliability of the information sources. It follows then that if one holds these two factors constant between information sets and lets all else (i.e., coherence) vary, then any effects of differences in coherence between the sets will be detected.

B&H's impossibility result for Bayesian Coherentism seeks to show that coherence can only be given a probabilistic, complete and transitive ordering relation if it is not separable (i.e., *separability* (BC_1) is inconsistent with *probabilism* and *ordering* (BC_2)). B&H intend their result to apply to any putative ordering, and thus to any proposed order-inducing probabilistic measure of coherence. They write, "Our strategy will be to show that *any* coherence measure would leave (BC_1) and (BC_2) vulnerable to counter-examples. Hence, no reasonable proposal for a coherence measure could ever succeed" (2003, 20).

In presenting such counter-examples, B&H utilize their own reformulation of Bayes's theorem:[4]

$$P(R_1, \ldots, R_n|REPR_1, \ldots, REPR_n) = P^*(R_1, \ldots, R_n) = \frac{a_0}{a_0 + a_1\bar{r} + \ldots + a_n\bar{r}^n}$$

This version of Bayes's theorem presents the posterior probability of an information set entirely as a function of the reliability of the information sources[5] and

[2] B&H clarify that "independence" here refers to the fact that "R_i *screens off* $REPR_i$ from all other fact variables R_j and from all other report variables $REPR_j$" (2003, 16).

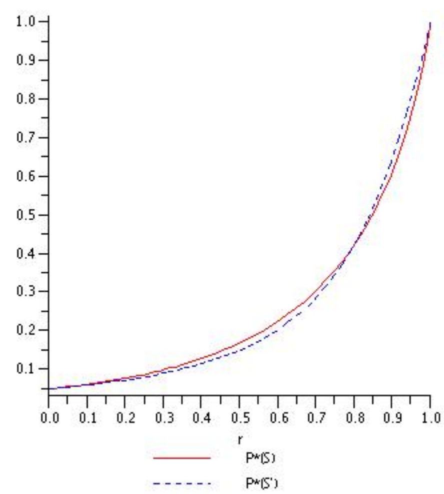
[3] In this paper, I use italics to denote the value of a variable and non-italics to denote variables; e.g., the variable R_1 can take on two values, R_1 and $\neg R_1$.

[4] See B&H (2003)'s appendix A.1 for the derivation of this form of Bayes's theorem.

[5] There is an important distinction to be made here: B&H assume *for simplicity* that all information sources are equally reliable *within* an information set, and - as already noted - they assume *for ceteris paribus' sake* that information sources are equally reliable across information sets; thus, r is only given one general value in these examples, which represents the level of reliability of all information sources for all information sets.

what B&H call the "weight vector" of an information set. Any information set has a corresponding weight vector $<a_0, a_1, ...a_n>$ where each element a_i of the weight vector represents "the sum of the joint probabilities of all combinations of i negative values and $n-i$ positive values of the variables R_1, ..., R_n" (2003, 17).[6]

B&H proceed to give their putative counter-examples to Bayesian Coherentism by assuming (BC_2) and introducing certain information sets - defined by their weight vectors and thus by their probabilistic information - that obey separability's ceteris paribus conditions but nonetheless do not obey separability (BC_1). As an example, B&H put forth information sets $S = \{R_1, R_2, R_3\}$ and $S' = \{R'_1, R'_2, R'_3\}$ with respective weight vectors $< a_0, a_1, a_2, a_3 > = < .05, .30, .10, .55 >$ and $< a'_0, a'_1, a'_2, a'_3 > = < .05, .20, .70, .05 >$. B&H stipulate that the reliability of the information sources for these sets is equal and that the information sources are independent for both sets; thus, given that $a_0 = a'_0$, (BC_1)'s ceteris paribus conditions are enforced. Nonetheless, if one calculates $P^*(R_1, \ldots, R_n)$ and $P^*(R'_1, \ldots, R'_n)$ and allows the value of r to range from 0 to 1, then *regardless of the coherence measure that one prefers or the result that such a measure gives*, (BC_1) does not hold true. If our chosen coherence measure gives the result that $coh(S') \geq coh(S)$, then (BC_1) is violated for any value of $r \in (.8, 1)$ since $P^*(R_1{}', R'_2, R'_3) < P^*(R_1, R_2, R_3)$ in this interval. On the other hand, if our measure tells us that $coh(S') < coh(S)$, then (BC_1) is violated for any value of $r \in (0, .8]$ given that $P^*(R_1{}', R'_2, R'_3) \geq P^*(R_1, R_2, R_3)$ in this interval. This result is captured visually by the "criss-crossing effect" in the following graph:



[6]B&H offer the following example: "for an information triple containing the propositions R_1, R_2, and R_3, $a_2 = P(\neg R_1, \neg R_2, R_3) + P(\neg R_1, R_2, \neg R_3) + P(R_1, \neg R_2, \neg R_3)$. That is, a_2 is the sum of the joint probabilities of all combinations with two negative values and one positive value." Note that a_0 will always equal what B&H call the "expectedness of the information" for any information set.

Both of these results hold even in spite of the fact that B&H's ceteris paribus conditions are enforced. Thus, B&H conclude, (BC_1) and (BC_2) cannot be true together; they are inconsistent.[7]

The Problem with B&H's Ceteris Paribus Conditions

The efficacy of B&H's impossibility result depends on one's choice of ceteris paribus conditions. Thus, B&H (2006) show that their impossibility result works (i.e., there exist counter-examples to Bayesian Coherentism) for $n \geq 2$ if one only holds constant the reliability of the information sources (leaving expectedness of the information out of the ceteris paribus conditions); however, adding expectedness of the information to the ceteris paribus conditions, there exist counter-examples only for $n \geq 3$. This fact suggests that one might be able to sidestep the impossibility result simply by adding more ceteris paribus conditions. Indeed, B&H (2003, 21) write:

> [W]e have shown that our degree of confidence is a function of the reliability r and the weight vector $<a_0, ..., a_n>$. It may well be the case that there is another determinant D of our degree of confidence which differs from reliability, expectance, and coherence and which is also a function of r and $<a_0, ..., a_n>$. (BC_1) may well be true [i.e., the impossibility result may not hold] if we keep the reliability, the expectance, *as well as D* fixed under the ceteris paribus clause.

Here I show that a similar but different end-around the impossibility result is possible. The ceteris paribus conditions that B&H enforce have a certain intuitive appeal to them; it *does* seem right in the testimonial context to distinguish coherence from reliability and expectedness of the information. However, depending upon the specific measure of coherence that one considers, these conditions - and particularly the expectedness of information condition - might make very little sense. The choice of coherence measure intersects with one's choice of ceteris paribus conditions; such conditions may very well change depending on the measure one has in mind. In this way, I argue that the suitability of B&H's ceteris paribus conditions is not so clear-cut as they seem to think.

I show this using the following coherence measure as a case study: Tomoji Shogenji (1999) has proposed a measure of coherence according to which the degree of coherence of an information set corresponds to the degree to which the members of that set are relevant to one another as measured by the ratio:

$$C(\{R_1, ..., R_n\}) =_{def} \frac{P(R_1, ... R_n)}{P(R_1) \times ... \times P(R_n)}$$

Examining this measure, if an information set is composed of information items, all of which are statistically independent of one another piecewise and

[7] Note that this counter-example in and of itself only shows that (BC_1) and (BC_2) are inconsistent for information triples. B&H note this and proceed to show that their result extends to information sets where $n > 3$.

collectively, then the numerator will equal the denominator and thus $C(\{R_1, ...R_n\}) = 1$. If there is any positive relevance among the information items, $C(\{R_1, ...R_n\}) > 1$; and in the case that the information items are negatively relevant, $C(\{R_1, ...R_n\}) < 1$.

Shogenji endorses and argues for a particular ceteris paribus condition that is distinct from any one of B&H's conditions. He suggests that the denominator term of his measure $P(R_1) \times ... \times P(R_n)$ represents the degree of "total individual strength" of an information set. More specifically, according to Shogenji, "The values of these denominators depend on the number and specificities of the individual beliefs - the more beliefs the set contains and the more specific each belief is, the lower the value is." Upon the assumption that this total individual strength is wholly distinct from coherence, Shogenji proceeds to argue that one ought to include it in the ceteris paribus conditions necessary for observing the effects of differing levels of coherence (1999, 342):

> The impact of the beliefs' total individual strength on their truth indicates that we cannot evaluate truth conduciveness of coherence simply by checking whether more coherent beliefs are more likely to be true together than less coherent beliefs. Such comparison may lump together the effects of two factors - coherence and total individual strength - on truth. In order to evaluate truth conduciveness of coherence in isolation from the confounding factor, we must compare two sets that have the same total individual strength.

Consequently, in comparing the effects of different levels of coherence between information sets, Shogenji suggests that we hold the denominator of his measure constant between sets via our ceteris paribus conditions and allow for any difference to show up in the numerator. Although Shogenji (1999) doesn't explicitly discuss the testimonial context, it seems reasonable to assume that in this context he would endorse including the reliability of information sources in his ceteris paribus conditions as well. Thus, contra B&H, Shogenji endorses the following set of factors that influence our degrees of confidence:

1. Total individual strength: $P(R_1) \times ... \times P(R_n)$
2. Reliability of the information sources.
3. Coherence of the information.

Consequently, according to this taxonomy of epistemically-relevant factors, one ought to include (1) and (2) in her ceteris paribus conditions in order to observe the effects across information sets of differences in (3).

It is easy to see that B&H's ceteris paribus conditions make little sense in respect to Shogenji's measure. The term in the numerator of Shogenji's measure is exactly what B&H call the expectedness of information (a_0). Thus, B&H argue that this term should be held constant across sets as part of one's ceteris paribus conditions. Of course, if Shogenji were to adopt this additional condition, his measure would consequently be impotent in the detection of relative coherence between sets. In fact, the numerators of his measure would be required to be constant across sets by B&H's conditions and the denominators would be required to be held constant by Shogenji's own argument and all such information sets would consequently take the same measure of coherence. That is, these ceteris paribus conditions would restrain Shogenji's measure to apply only to sets that it renders equally coherent. Clearly then, Shogenji's measure - taken along with Shogenji's own ceteris paribus condition - is no fan of B&H's "expectedness of information" condition.

A Possibility Result for Bayesian Coherentism

At least two important questions follow in the wake of the previous section:

1. Given that B&H and Shogenji offer two different sets of ceteris paribus conditions, which set - if either - is appropriate for the screening off the effects of all epistemically relevant factors but coherence?

2. Can a set of ceteris paribus conditions different from B&H's (e.g., Shogenji's) allow the Bayesian Coherentist to sidestep the impossibility result?

While I make no attempt in this paper to answer the first question, I give an answer to the second question in this section. An extension of Shogenji's ceteris paribus conditions entails the consistency of Bayesian Coherentism - (BC_1) and (BC_2) are true together. This extension is not ad hoc; indeed, the ceteris paribus condition that it adds has - I argue - the same intuitive merits as the conditions for which Shogenji and B&H argue.

$P(REPR_1, ..., REPR_n)$ represents the expectedness of receiving our n reports from our information sources. This factor is epistemically relevant: all else being equal, as the expectedness of the reporting goes up, the level of confidence that we place in our information ought to go down (i.e., if we receive reports that we were expecting to receive anyway, then we will raise our confidence in the information being reported to a lesser degree than if we were to receive reports that were unexpected) and vice versa. Intuitively, this factor is separate from coherence; coherence - being a virtue of information sets *purely on the information level* - should not be affected by considerations having to do with the reporting of that information (including reliability and independence of the sources and expectedness of the reports). Adding this factor to Shogenji's taxonomy results in the following list of epistemically relevant factors in the testimonial context:

1. Total individual strength: $P(R_1) \times ... \times P(R_n)$

2. Expectedness of the reports: $P(REPR_1, ..., REPR_n)$

3. Reliability of the information sources.

4. Coherence of the information.

Given this setup, in order to meet separability's ceteris paribus conditions, one must hold factors (1), (2), and (3) constant between sets.

It is straightforward to show that if one enforces this set of ceteris paribus conditions, there can be no counter-examples to (BC_1) and (BC_2). Consider the simple form of Bayes's theorem:

$$P^*(R_1, \ldots, R_n) = \frac{P(R_1, ..., R_n) \times P(REPR_1, \ldots, REPR_n | R_1, \ldots, R_n)}{P(REPR_1, \ldots, REPR_n)}$$

Given that it is a general prerequisite of B&H's discussion of Bayesian Coherentism that information sources are independent - in the sense that "R_i *screens off* $REPR_i$ from all other fact variables R_j and from all other report variables $REPR_j$" - in comparing the posterior probabilities of two information sets, this can be rewritten:

$$P^*(R_1, \ldots, R_n) = \frac{P(R_1, ..., R_n) \times P(REPR_1 | R_1) \times \ldots \times P(REPR_n | R_n)}{P(REPR_1, \ldots, REPR_n)}$$

In this equation, all of the individual likelihood terms represent the true positive reporting rates for each of the n information sources. These terms are equal across sets as part of ceteris paribus condition (3) - reliability of the information sources.[8] Thus, in comparing the effects of differing coherence on posterior probability *ceteris paribus*, it will not be necessary to consider these terms. We need only compare:

$$\frac{P(R_1, ..., R_n)}{P(REPR_1, \ldots, REPR_n)}$$

The denominator term in this relation is simply the additional "expectedness of the reports" ceteris paribus condition. This term will be equal across sets in the ceteris paribus context, so we are left to compare the single term, $P(R_1, ..., R_n)$. Thus, enforcing this set of ceteris paribus conditions, relative values of posterior probability will be directly proportional to those of the expectedness of the evidence.

[8] One might respond with the insight that B&H's reliability ceteris paribus condition does not entail the equality of true positives between sets - i.e., B&H's reliability condition says that one should only compare sets for which $r := 1 - P(REPR_i | \neg R_i)/P(REPR_i | R_i)$ is equal; however the ratio $P(REPR_i | \neg R_i)/P(REPR_i | R_i)$ could of course be equal between sets even if the true positives $P(REPR_i | R_i)$ were unequal. Nonetheless, B&H do seem to require ultimately that not only must the ratios be equal across sets but also the true and false positives themselves: "To keep things simple, let us assume that all witnesses are equally reliable, i.e., $p_i = p$ and $q_i = q$ for all $i = 1, ..., n$" (B&H, 15). Additionally, it just seems to make good intuitive sense to hold true positive rates (and false positive rates) constant in our ceteris paribus conditions; indeed, Shogenji (2006) even argues that the true positive rate on its own could be taken to be a proper measure of witness reliability. Regardless, the above response is not necessarily relevant anyway as it is no longer B&H's specific ceteris paribus conditions that we need to have in mind at this point.

This is an appealing conclusion from Shogenji's point of view. The effect of different values of coherence ceteris paribus is reflected entirely in the value of $P(R_1, ..., R_n)$ in the testimonial context. As I suggested earlier, this is also true of Shogenji's measure in the non-testimonial context: given Shogenji's measure, an information set S is more coherent than another information set S' if and only if:

$$C(S) = \frac{P(R_1,...R_n)}{P(R_1)\times...\times P(R_n)} > \frac{P(R'_1,...R'_n)}{P(R'_1)\times...\times P(R'_n)} = C(S')$$

Relative levels of coherence between S and S' *ceteris paribus* - more specifically, given condition (1) above - will thus directly correspond to the relation between $P(R_1, ..., R_n)$ and $P(R'_1, ..., R'_n)$. That is, when Shogenji's "total individual strength" condition is enforced and his measure is used, differences in coherence are directly proportional to differences in the expectedness of the information. Given that our ceteris paribus conditions lead to the result above - that relative values of posterior probability $P^*(R_1, \ldots, R_n)$ are directly proportional to those of the expectedness of the information $P(R_1, \ldots, R_n)$ in the testimonial context - we get the following for all information sets S, S'∈***S***:

- $C(S) > C(S')$ *iff* $P^*(R_1, \ldots, R_n) > P^*(R'_1, \ldots, R'_n)$
- $C(S) < C(S')$ *iff* $P^*(R_1, \ldots, R_n) < P^*(R'_1, \ldots, R'_n)$
- $C(S) = C(S')$ *iff* $P^*(R_1, \ldots, R_n) = P^*(R'_1, \ldots, R'_n)$

These results entail (BC_1). Thus, assuming (BC_2) and enforcing the extended set of ceteris paribus conditions, (BC_1) follows. Given this particular set of ceteris paribus conditions, *separability* is consistent with *probabilism* and *ordering*.

Concluding Remarks

This paper makes no attempt to argue for any particular set of ceteris paribus conditions for coherence. Rather, I have merely shown (using Shogenji's work as a case study) that there exist sets - at least one - of ceteris paribus conditions that are intuitively appealing and do allow one to avoid B&H's impossibility result. B&H's result attempts to show that Bayesian Coherentism is an impossible position given that its fundamental tenets are inconsistent, but I have shown that their result relies on a presumed set of ceteris paribus conditions that can be rejected. Manifestly and crucially, the set of ceteris paribus conditions that one needs to enforce in the testing for separability will change depending upon the putative measure of coherence that he adopts. Thus, B&H's mistake is to rely on a specific set of ceteris paribus conditions in deriving their impossibility result, which in turn purports to apply across the board to all proposed measures of coherence. One simply cannot stipulate such conditions that will be appropriate to all such measures. If one changes the measure being considered, she may very well have to change the ceteris paribus conditions considered.

B&H could of course respond by noting that while different coherence measures may call for different conditions, there is only ultimately one set of ceteris paribus conditions that is truly appropriate for the testing of the separability of coherence. I would agree. However, this observation shifts the debate from the merits of B&H's result to the suitability of their ceteris paribus conditions. That is, before their impossibility result can be deemed successful or unsuccessful, B&H need to convince us that their ceteris paribus conditions - as opposed to other seemingly plausible options - are the appropriate ones. Such convincing must stretch beyond the typical intuition-based taxonomies of epistemically relevant factors. Very little to no argument is actually given by B&H [(2003), (2005), (2006)] in this regard.

References

[1] Bovens, L. and Hartmann, S. (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.

[2] Bovens, L. and Hartmann, S. (2005). "Why There Cannot be a Single Probabilistic Measure of Coherence." *Erkenntnis* 63: 361-74.

[3] Bovens, L. and Hartmann, S. (2006). "An Impossibility Result for Coherence Rankings." *Philosophical Studies* 128: 77-91.

[4] Fitelson, Branden. (2003). "A Probabilistic Theory of Coherence." *Analysis* 63.3: 194-99.

[5] Fitelson, Branden. (2004). "Two Technical Corrections to My Coherence Measure." Unpublished.

[6] Olsson, Erik J. (2005). *Against Coherence*. Oxford: Oxford University Press.

[7] Shogenji, T. (1999). "Is Coherence Truth Conducive?" *Analysis* 59.4: 338-45.

[8] Shogenji, T. (2006). Review of Luc Bovens and Stephan Hartmann, *Bayesian Epistemology*. *Theoria* 72: 166-71.