

On the alleged impossibility of Bayesian Coherentism

Jonah N. Schupbach

Published online: 9 October 2007
© Springer Science+Business Media B.V. 2007

Abstract The success of Bovens and Hartmann’s recent “impossibility result” against Bayesian Coherentism relies upon the adoption of a specific set of *ceteris paribus* conditions. In this paper, I argue that these conditions are not clearly appropriate; certain proposed coherence measures motivate different such conditions and also call for the rejection of at least one of Bovens and Hartmann’s conditions. I show that there exist sets of intuitively plausible *ceteris paribus* conditions that allow one to sidestep the impossibility result. This shifts the debate from the merits of the impossibility result itself to the underlying choice of *ceteris paribus* conditions.

Keywords Bayesianism · Bayesian Coherentism · Coherence · Impossibility result · *Ceteris paribus* conditions · Coherence measures

1 B&H on Bayesian Coherentism

In several recent publications (2003, 2005, 2006), Luc Bovens and Stephan Hartmann (hereon, “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, pp. 78–9, 2003, pp. 11–12, 25):

- *Separability* (BC_1): For all information sets $S, S' \in \mathcal{S}$, if S is no less coherent than 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...” over \mathcal{S} is fully determined by the probabilistic features of the information sets contained in \mathcal{S} .

J. N. Schupbach (✉)
History & Philosophy of Science, University of Pittsburgh, 1017 Cathedral of Learning, Pittsburgh,
PA 15260, USA
e-mail: jns24@pitt.edu

- *Ordering* (BC_{2[iii]}): The binary relation of “...being no less coherent than...” is an ordering; i.e., the relation is transitive and complete.¹

Importantly, B&H propose specific, intuitively plausible *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 our intuitions tell us that there are three (and only three) distinct epistemically relevant “factors that affect our degree of confidence:”

1. Expectedness of the information: i.e., joint probability $P(R_1, R_2, \dots, R_n)$ where R_i represents the i 'th “information item” in the information set.³
2. Reliability of the information sources: $r = 1 - q/p$ where $q = P(\text{REPR}_i | \neg R_i)$ and $p = P(\text{REPR}_i | R_i)$. 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 intuition-based taxonomy of epistemically relevant factors, B&H propose that in order to meet *separability*'s *ceteris paribus* conditions, one must hold factors (1) and (2) equal between sets. That is, in order to detect the effects of a difference in coherence per se between sets, one needs to hold all else equal, and this “all else,” per B&H, consists of the expectedness of the information and the reliability of the information sources.

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* (BC₁) is inconsistent with *probabilism* and *ordering* (BC₂)). 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 (BC₁) and (BC₂) vulnerable to counter-examples. Hence, no reasonable proposal for a coherence measure could ever succeed” (2003, p. 20).

In order to present such counter-examples, B&H reformulate Bayes's theorem via their “independence of information sources” assumption (where $\bar{r} = 1 - r$):⁴

$$P(R_1, \dots, R_n \mid \text{REPR}_1, \dots, \text{REPR}_n) = P^*(R_1, \dots, R_n) = \frac{a_0}{\sum_{i=0}^n a_i \bar{r}^i}$$

This variant of Bayes's theorem presents the posterior probability of an information set entirely as a function of the reliability of the information sources⁵ and what B&H

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

² B&H require witnesses to be 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 ” (2003, 16).

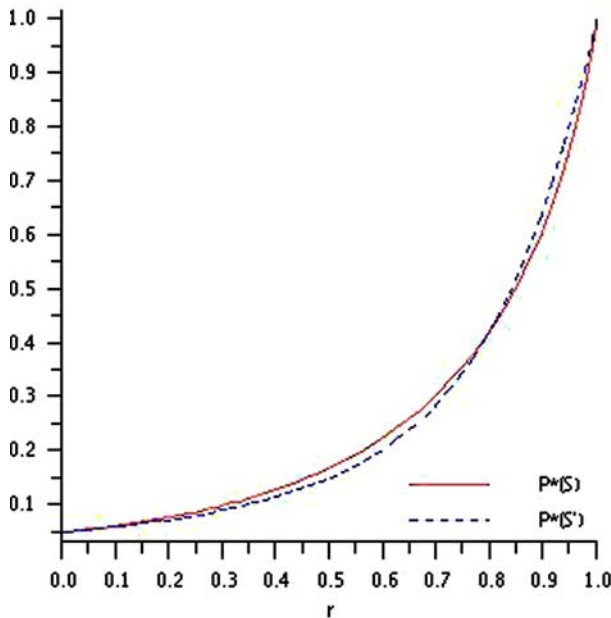
³ In this paper, I follow B&H in using italics to denote variables and non-italics to denote the values of variables; e.g., the variable R_i can take on two values, R_i and $\neg R_i$. Also following B&H, commas denote conjunction; thus, $P(R_1, \dots, R_n) = P(R_1 \& \dots \& R_n)$.

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 items across all information sets in question.

call the “weight vector” of an information set. Any information set has a corresponding weight vector $\langle a_0, a_1, \dots, a_n \rangle$ 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, \dots, R_n ” (2003, p. 17).⁶

B&H develop their attempted counter-examples to Bayesian Coherentism by assuming (BC₂) 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* (BC₁). As one such 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 $\langle a_0, a_1, a_2, a_3 \rangle = \langle .05, .30, .10, .55 \rangle$ and $\langle a_0', a_1', a_2', a_3' \rangle = \langle .05, .20, .70, .05 \rangle$. B&H stipulate that the reliability of the information sources for these sets is equal; thus, given that $a_0 = a_0'$, (BC₁)'s ceteris paribus conditions are enforced. Nonetheless, if one calculates $P^*(R_1, R_2, R_3)$ and $P^*(R_1', R_2', R_3')$ and allows the value of r to range from 0 to 1, then *regardless of the coherence measure that one prefers and the result that such a measure gives*, (BC₁) does not hold true. If our chosen coherence measure gives the result that $\text{coh}(S') \geq \text{coh}(S)$, then (BC₁) 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 $\text{coh}(S') < \text{coh}(S)$, then (BC₁) 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, and the sum of the elements of any weight vector is always equal to one.

Both of these results hold even in spite of the fact that B&H's *ceteris paribus* conditions are enforced. Thus, B&H conclude that if one assumes (BC₂), then there exist cases where a set is more coherent but less probable than another set *ceteris paribus*. Consequently, (BC₁) and (BC₂) cannot be true together; they are inconsistent.⁷

2 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, p. 21) write:

[W]e have shown that our degree of confidence is a function of the reliability r and the weight vector $\langle a_0, \dots, a_n \rangle$. 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 $\langle a_0, \dots, a_n \rangle$. (BC₁) 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 ultimately different response to 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 one might think.

I show this using the following case study: 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:

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

⁷ This counter-example, in and of itself, only shows that (BC₁) and (BC₂) are inconsistent for information triples. B&H note this and proceed to show that their result extends more generally to information sets where $n > 3$.

Examining this measure, if an information set is composed of mutually independent pieces of information, then the numerator will equal the denominator and thus $C(R_1, \dots, R_n) = 1$. If there is any positive relevance among the information items, $C(R_1, \dots, R_n) > 1$; and in the case that the information items are negatively relevant, $C(R_1, \dots, R_n) < 1$.

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 $P(R_1) \times \dots \times P(R_n)$ 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 intuitively plausible assumption that this total individual strength is wholly distinct from coherence, Shogenji (1999, p. 342) proceeds to argue that one ought to include it in the *ceteris paribus* conditions necessary for observing the effects of differing levels of coherence:

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) does not 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; in this case, the following set of factors influence our degrees of confidence:

1. Total individual strength: $P(R_1) \times \dots \times P(R_n)$.
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) as *ceteris paribus* conditions in order to observe the effects across information sets of differences purely in (3).

It is easy to see that B&H’s *ceteris paribus* conditions make little sense if one endorses Shogenji’s measure. The term in the numerator of Shogenji’s measure is that which B&H call the expectedness of information (a_0). As previously noted, B&H assert that this term should be held constant across sets as part of one’s *ceteris paribus* conditions. But if Shogenji were to adopt this additional condition, his measure would consequently be impotent in the detection of the effects of coherence per se between sets. In fact, the numerators of Shogenji’s 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; hence, all such information sets would 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, for one who accepts Shogenji's measure, it makes no sense to include expectedness of the information in the *ceteris paribus* conditions; indeed, this is exactly the term that such a person would expect to vary in value with differing values of coherence *ceteris paribus*.

3 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 intuitively plausible 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 some intuitively plausible 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—at least according to my intuitions—the same intuitive merits as the conditions for which Shogenji and B&H argue.

$P(\text{REPR}_1, \dots, \text{REPR}_n)$ represents the expectedness of receiving our n reports from our information sources. This factor is epistemically relevant: according to Bayes's theorem, all else being equal, as the expectedness of the reporting increases, the level of confidence that we place in our information upon receiving these reports ought to decrease (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 distinct 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 of the sources and expectedness of the reports). Indeed, it is easy to imagine cases in which we very much expect to receive reports of incoherent information or in which we are doubtful that we will receive reports of very coherent information. Thus, our expectedness of the reports intuitively seems distinct from coherence. Following these intuitions, one might easily be lead to add this factor to Shogenji's taxonomy resulting in the following list of epistemically relevant factors in the testimonial context:

1. Total individual strength: $P(R_1) \times \dots \times P(R_n)$.
2. Expectedness of the reports: $P(\text{REPR}_1, \dots, \text{REPR}_n)$.
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 (BC₁) and (BC₂). Consider the simple form of Bayes's theorem:

$$P^*(R_1, \dots, R_n) = \frac{P(R_1, \dots, R_n) \times P(\text{REPR}_1, \dots, \text{REPR}_n | R_1, \dots, R_n)}{P(\text{REPR}_1, \dots, \text{REPR}_n)}$$

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*”—this can be rewritten:

$$P^*(R_1, \dots, R_n) = \frac{P(R_1, \dots, R_n) \times P(\text{REPR}_1 | R_1) \times \dots \times P(\text{REPR}_n | R_n)}{P(\text{REPR}_1, \dots, \text{REPR}_n)}$$

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 the reliability ceteris paribus condition.⁸ Thus, in comparing the effects of differing levels of coherence on posterior probability ceteris paribus, it is not necessary to consider these terms. We need only compare:

$$\frac{P(R_1, \dots, R_n)}{P(\text{REPR}_1, \dots, \text{REPR}_n)}$$

The denominator term in this ratio 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(R₁, ..., 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 information.*

Similarly, given that the ceteris paribus conditions include Shogenji's total individual strength condition, it is also the case that ceteris paribus:

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

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 $r := 1 - P(\text{REPR}_i | \neg R_i) / P(\text{REPR}_i | R_i)$ is equal; however the ratio $P(\text{REPR}_i | \neg R_i) / P(\text{REPR}_i | R_i)$ could of course be equal between sets even if the true positives $P(\text{REPR}_i | R_i)$ 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 $i = 1, \dots, n$ ” (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.

if and only if $P(R_1, \dots, R_n) > P(R'_1, \dots, R'_n)$. More generally, *enforcing this set of ceteris paribus conditions, relative values of coherence between information sets will be directly proportional to those of the expectedness of the information.* Consequently (given that if relative values of two different terms are directly proportional to relative values of the same third term, then they are directly proportional to each other), ceteris paribus, different levels of coherence are directly proportional to those of the posterior probability of the information. This is just a restatement of (BC_1) . Therefore, for one who accepts Shogenji's measure and the extended—though still intuitively plausible—set of ceteris paribus conditions, (BC_1) follows. Accordingly, in this case, *separability* (BC_1) is consistent with *probabilism* and *ordering* (BC_2) .

4 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 or she 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, he or 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.⁹

⁹ Thanks very much to Luc Bovens, Clark Glymour, Stephan Hartmann, Tomoji Shogenji, Greg Wheeler, and especially to Adam True Bendorf for helpful correspondence.

References

- Bovens, L., & Hartmann, S. (2006). An impossibility result for coherence rankings. *Philosophical Studies*, 128, 77–91.
- Bovens, L., & Hartmann, S. (2005). Why there cannot be a single probabilistic measure of coherence. *Erkenntnis*, 63, 361–374.
- Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.
- Shogenji, T. (2006). Review of Luc Bovens and Stephan Hartmann, *Bayesian epistemology*. *Theoria*, 72, 166–171.
- Shogenji T. (1999). Is coherence truth conducive? *Analysis*, 59(4), 338–345.