

A Corrective to Bovens and Hartmann's Measure of Coherence

Wouter Meijls

Abstract

Bovens and Hartmann (2003a, 2003b) propose to analyze coherence as a confidence-boosting property. On the basis of this idea, they construct a new probabilistic theory of coherence. In this paper, I will attempt to show that the resulting measure of coherence clashes with some of the intuitions that motivate it. Also, I will try to show that this clash is not due to the view on coherence as a confidence-boosting property or to the general features of the model that Bovens and Hartmann use to analyze coherence. It will turn out that there is at least one other measure that is similarly based on the concept of a confidence-boosting property, but does not have the same counterintuitive results.

1 Introduction

In their *Bayesian Epistemology* (2003a), Luc Bovens and Stephan Hartmann persuasively argue that one cannot define a measure of coherence without making explicit what role coherence is to play. According to them, it makes sense to talk of the coherence of a law firm or an ant-hill, but it makes no sense to argue that they serve the same purpose: coherence in a law firm makes it more efficient, in an ant-hill it aids the survival of the colony (*op. cit.*: 31). Consequently, any attempt to construct a *measure* of coherence should first indicate what type of coherence one wishes to measure, and what role coherence is supposed to play. According to Bovens and Hartmann, the philosophically interesting notion of coherence is as a property of an information set with the role of boosting our confidence in that information set (*ibid.*). On the basis of this idea, they construct a probabilistic theory of coherence.

The result of their approach is not an actual *measure* of coherence, in the sense that each information set is assigned a specific value of coherence. Instead, they define the relation of *being no less coherent than*, ' \succcurlyeq '. As they explain (*op. cit.*: 3, 35), this relation induces an ordering on the set of all information sets which is not complete, meaning that it will not necessarily be the case that, given two sets of propositions, one of the two is the most coherent. It may be arguable that this indeterminacy is not problematic when it occurs in a case in which it is intuitively not clear which of two sets is the more coherent. However, Douven and Meijls (2004) and Meijls and Douven (2004) have shown that sometimes Bovens and Hartmann's theory of coherence will also refrain from judgement if one of the two sets is intuitively clearly more coherent than the other. Hence in some cases their measure behaves in an intuitively unsatisfying way. However, the case is even worse: in section 3 of this paper I will show that their measure also fails to satisfy two very general intuitions about coherence.

Nevertheless, the overall goal of this paper is constructive rather than destructive, for in section 4 I show that we need not blame Bovens and Hartmann's general approach for these counterexamples. Instead, the problem may lie in an additional

element, which I will call the maximality requirement, that they add to their theory. In section 5, I will show how abandoning the maximality requirement will lead to a new measure of coherence that is equally based on coherence as a confidence-boosting property, but that does not yield the same counterintuitive consequences. The conclusion of this paper, therefore, will be that the examples proposed in this and earlier papers do not necessarily challenge Bovens and Hartmann's basic intuitions. I will start by briefly introducing Bovens and Hartmann's theory of coherence.

2 Bovens and Hartmann's Theory of Coherence

Define the notion of an *information set* S as a finite set of propositions, about each of which we have been informed by a separate source. Call these information sources *witnesses* and assume that they are independent of each other and neither fully reliable nor fully unreliable. Also, suppose that they supply only positive or negative reports, that is, the witnesses report that something is or is not the case; they do not report probabilities. Let \mathbf{S} be the set of all such information sets.

Assuming positive reports on all of the propositions in an information set, what are the factors that determine our confidence that the conjunction of the propositions in that set is true? According to Bovens and Hartmann, there are three such factors. These are, respectively, the prior probability that the information is true; the reliability of the witnesses; and the measure in which the propositions cohere. These factors should be expressed as *ceteris paribus* claims. The *ceteris paribus* condition for coherence reads: '[w]hen we gather information, then the more coherent the story that materializes is, the more confident we may be, *ceteris paribus*' (2003a: 11). Similarly, an increase either in the prior probability of the conjunction of propositions in an information set S or in the witness reliability will make us more confident that the information is true, *ceteris paribus*.

In order to explain how Bovens and Hartmann determine the probability that the conjunction of propositions in a set S is true, given positive reports on all of these propositions, it is necessary to introduce some terminology. Given an information set $S = \{R_1, \dots, R_n\}$ and a probability distribution $p(\cdot)$ over the elements of S , let ' a_i ' stand for the sum of the probabilities of the conjunctions consisting of $n - i$ elements of S and the negations of all the remaining elements of S . So, for instance, if $S = \{R_1, R_2\}$, then $a_0 = p(R_1 \wedge R_2)$, $a_1 = p(R_1 \wedge \neg R_2) + p(\neg R_1 \wedge R_2)$, and $a_2 = p(\neg R_1 \wedge \neg R_2)$. Furthermore, let REPR_j be a report by witness j to the effect that R_j is the case (there is one witness for each proposition). To model the reliability of the witnesses, let

$$\mathsf{p}_j =_{df} p(\text{REPR}_j \mid R_j) \quad (1)$$

be the true-positive rate of witness j with respect to proposition R_j and let

$$\mathsf{q}_j =_{df} p(\text{REPR}_j \mid \neg R_j) \quad (2)$$

be the rate of false positives. Assume that all witnesses are equally reliable: for all j , we have $\mathsf{p}_j = \mathsf{p}$ and $\mathsf{q}_j = \mathsf{q}$. Next, define

$$r =_{df} 1 - \mathsf{q}/\mathsf{p} \quad (3)$$

as the reliability of the witnesses. They are assumed to be neither fully reliable ($r \neq 1$) nor fully unreliable ($r \neq 0$); thus, $r \in (0, 1)$.

With this terminology in place, let us look at Bovens and Hartmann's derivation of their measure of coherence. They start with calculating the posterior probability

$$p^*(R_1 \wedge \cdots \wedge R_n) =_{df} p(R_1 \wedge \cdots \wedge R_n | \text{REPR}_1 \wedge \cdots \wedge \text{REPR}_n), \quad (4)$$

which gives the probability of the conjunction of the elements of set S given positive reports with respect to all of these elements. Now it can be shown (see Bovens and Hartmann 2003a, appendix A1) that if $a_0 > 0$, the following relation holds:

$$p^*(R_1 \wedge \cdots \wedge R_n) = \frac{a_0}{\sum_{i=0}^n a_i (1-r)^i}. \quad (5)$$

The *probability boost* $b(R_1 \wedge \cdots \wedge R_n)$ is defined as the ratio of the posterior probability and the prior probability:

$$b(R_1 \wedge \cdots \wedge R_n) = \frac{p^*(R_1 \wedge \cdots \wedge R_n)}{p(R_1 \wedge \cdots \wedge R_n)}. \quad (6)$$

As noted above, Bovens and Hartmann believe that there are three different factors that influence the probability boost: witness reliability, prior probability and coherence. To find a true measure of coherence, therefore, one needs to distinguish the degree to which coherence contributes to the value of $b(R_1 \wedge \cdots \wedge R_n)$ from the contributions of witness reliability and prior probability. But it is not straightforward how this may be achieved. The basic problem is that although reliability may be convincingly portrayed as being independent from coherence, this does not hold for the prior probability. To see this, remember that the prior probability of the information set is the prior probability that all the elements of the information set are true, or, in set-theoretic terms, the measure in which they overlap. It should be uncontroversial (and Bovens and Hartmann agree) that the measure in which the propositions overlap is very important for the measure in which they cohere. What this means is that the above *ceteris paribus* condition will be of little help for determining a measure of coherence. If the thing that plays an important role for coherence (to wit, the prior probability) must be kept constant for the definition to apply, then we will have no way of measuring the contribution of the prior probability to the coherence. An additional element or criterion is required to disentangle the prior probability and the coherence of a set.

The element that Bovens and Hartmann add is what I will call the *maximality requirement*. In their view a set is maximally coherent if it consists of equivalent propositions (*op. cit.*: 32–35), where two propositions R_i and R_j are equivalent if they are logically equivalent relative to the background knowledge.¹ Let such a set be an *equivalent set* and let a set be *nonequivalent* if not all of its propositions are equivalent relative to the background knowledge. Furthermore, let two sets of propositions S_i and S_j be equivalent if the conjunctions of the propositions in the sets are equivalent relative to the background knowledge. According to Bovens and Hartmann, the property of maximal coherence should not depend on a set's prior probability or the number of propositions. Using this requirement, they define the notion of a maximal confidence boost, i.e., the confidence boost a set would have received if the information set had consisted of equivalent propositions, in other words, if all of a_1, \dots, a_{n-1} had been zero (Bovens and Hartmann 2003a: 32–33). In that case $a_n = 1 - a_0$,² and so equation (5) reduces to

$$p_{\max}^*(R_1 \wedge \cdots \wedge R_n) = \frac{a_0}{a_0 + a_n (1-r)^n}. \quad (7)$$

¹ That is, if $K \wedge R_i \equiv K \wedge R_j$, where K is the background knowledge. Following Bovens and Hartmann, I will suppress reference to the background knowledge throughout this paper.

² Since $\sum_{i=0}^n a_i = 1$ and in this case $\sum_{i=1}^{n-1} a_i = 0$, we have $a_0 + a_n = 1$.

Next, one can define a measure $c_r(S)$ as the fraction of the actual boost over the maximum boost (for a derivation, see Bovens and Hartmann 2003a: 34):

$$c_r(S) =_{df} \frac{b(R_1 \wedge \dots \wedge R_n)}{b_{\max}(R_1 \wedge \dots \wedge R_n)} = \frac{a_0 + (1 - a_0)(1 - r)^n}{\sum_{i=0}^n a_i(1 - r)^i}. \quad (8)$$

It is easy to see that $c_r(S) = 1$ if R_1, \dots, R_n are equivalent, and $c_r(S) < 1$, otherwise. Next, the *difference function* calculates the difference between the values of c_r for two different sets:

$$f_r(S, S') = c_r(S) - c_r(S'). \quad (9)$$

Of course, this function still depends critically on the reliability of the witnesses. All else equal, a higher value for r will raise the value of $f_r(S, S')$. According to Bovens and Hartmann, coherence should be independent of witness reliability, and they cancel out the dependence on the reliability by defining the two-place relation of *being no less coherent than*, ‘ \succcurlyeq ’, as follows: for all $S, S' \in \mathbf{S}$,

$$S \succcurlyeq S' \text{ iff } f_r(S, S') \geq 0 \text{ for all values of } r \in (0, 1). \quad (10)$$

Interestingly, on this account there is no such thing as the coherence of a set: it is impossible to say whether a set is (in)coherent. Instead, only comparative judgements are possible: a set S is no less coherent than another set S' iff $f_r(S, S') \geq 0$ for all values of the reliability parameter.³ Also, since $f_r(S, S')$ may be positive for some values of r and negative for others, this analysis yields only a quasi-ordering: it is reflexive and transitive, but not complete. That is, there may be sets S and S' such that neither $S \succcurlyeq S'$ nor $S' \succcurlyeq S$. According to Bovens and Hartmann this is a desirable result, since in some cases we are intuitively in the dark with respect to the question which of two sets is the more coherent. However, as the next section will show, even in some cases where we are very sure about which is the more coherent of two sets, this analysis equally fails to yield a verdict.

3 Two Counterintuitive Consequences

In Meijs and Douven (2004) and Douven and Meijs (2004), three different examples of counterintuitive consequences of the theory of Bovens and Hartmann have been put forward. All of these proceed by presenting two sets of which it is intuitively very clear which one is more coherent, but of which Bovens and Hartmann’s theory withholds judgement. I will not rehearse these examples here. Instead, I will present two different examples, which show that the difference function fails to satisfy two general intuitions about coherence – intuitions, for that matter, that Bovens and Hartmann seem to endorse.

Independence and Positive Dependence. In his (2003), Branden Fitelson argues that independent sets should always be less coherent than positively dependent sets and more coherent than negatively dependent sets. A general definition of positive dependence requires that for all nonempty subsets S_i and S_j of information set S , if $S_i \cap S_j = \emptyset$, then $p(S_i | S_j) > p(S_i)$; for negative dependence the sign is reversed and for independence $p(S_i | S_j) = p(S_i)$ is used. Intuitively, it makes quite a lot of sense to require that independent sets are always less coherent than positively dependent sets.

³ One may feel that the introduction of the ‘ \succcurlyeq -relation as ‘being no less coherent than’ is rather awkward also. Would we not want our theory of coherence to produce a ‘ \succ ’ relation of being ‘more coherent than’? Nevertheless, this can easily be remedied by stipulating that $S \succ S'$ iff $f_r(S, S') \geq 0$ for all values of $r \in (0, 1)$ and $f_r(S, S') > 0$ for at least one value of $r \in (0, 1)$.

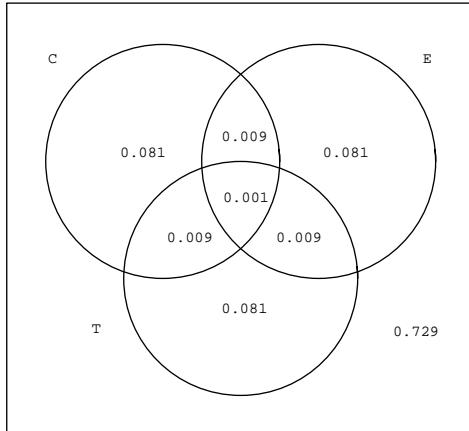


Fig. 1: A diagram for the probabilities of set S^I .

To illustrate this, consider the following example, originally due to BonJour (1985: 96):⁴

$$S^I = \{C = [\text{This chair is brown}], E = [\text{Electrons are negatively charged}], T = [\text{Today is Thursday}]\}.$$

In this set, the propositions may be assumed to be independent: the fact that it is Thursday does not seem to affect the colour of the chair or the charge of electrons, and vice versa. Also, we would feel hard pressed to call this set coherent given any specification of the probabilities that makes these propositions come out independent.

Moreover, Bovens and Hartmann agree with this general intuition. Following BonJour, they compare the first set with another of which ‘there is no doubt’ (Bovens and Hartmann 2003a: 29) that it is more coherent. I will give a slightly different version of that set, since in BonJour’s example, it is still the case that two of the three propositions are independent. Therefore, consider set S^P :

$$S^P = \{B = [\text{This bird is black}], R = [\text{This bird is a crow}], L = [\text{This bird has a life-long mate}]\}.$$

In this second set, the three propositions greatly support each other: in general crows are black and establish life-long relationships and vice versa. According to Fitelson’s intuition, this implies that set S^I must be more coherent than set S^P . Since this example is actually stronger than the original example by BonJour (in this case all of the propositions support each other), again there can be no doubt that the second set is more coherent than the first. This intuition seems independent of the marginal probabilities of the propositions in these sets: as long as they are independent in the first set and positively dependent in the second, it seems to matter little how probable they are.

Interestingly, however, when Bovens and Hartmann discuss how the difference function fares with respect to this example, they choose a probability model in which the unconditional probabilities of the propositions are (almost) equal. It is highly uncertain whether Bovens and Hartmann chose this probability model on purpose. According to them what is essential about what I have called set S^I is

⁴ The brackets around the sentences are used to refer to the sentences’ propositional contents.

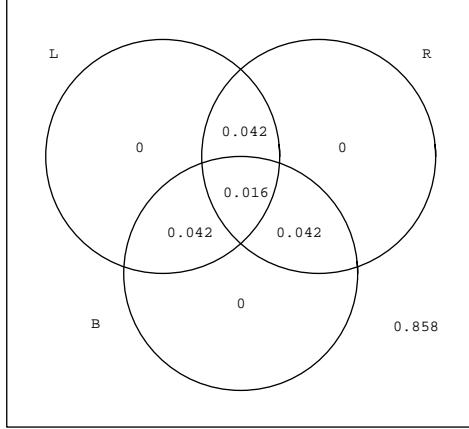


Fig. 2: A diagram for the probabilities of set S^P .

that it is probabilistically independent, not that it has almost the same values for the unconditional probabilities as the other set. Moreover, intuitively there seems to be little reason for any such restriction. It seems irrelevant for the question at hand whether or not the propositions in the first set are highly (im)probable. Instead, it appears that set S^P should be more coherent than S^I for all models in which the first is positively dependent and the second is independent.

Nevertheless, even if one accepts this additional constraint, there are still probability models in which it is not the case that S^P is more coherent than S^I . For example, consider a model for the sets S^I and S^P in which all six propositions have the same unconditional probability of .1. Since set S^I is independent, the value of $p(R_j^I)$ directly determines the values of a_0^I, \dots, a_3^I (one can read off these values from figure 1).

For set S^P assume that $a_2 = 0$, i.e., in the population we are studying (1) there are no crows that are neither black nor have a life-long mate, (2) there are no birds that have a life-long mate but are neither black nor crows and (3) there are no black birds that are neither crows nor have a life-long mate. Also, assume that the probability of a bird being a crow and black and having a life-long mate equals .016. Given the fact that the unconditional probabilities are all .1, the other probabilities follow directly (they can be read off from figure 2). It is not hard to see that in this case S^P is positively dependent and so, following Fitelson's suggestion, it should hold that S^P is more coherent than S^I .

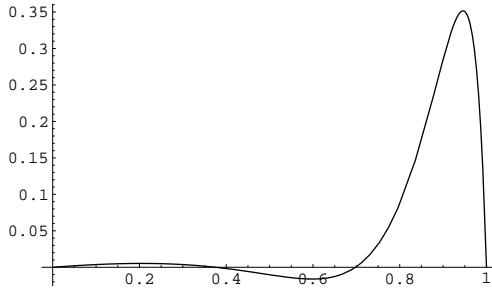
But now look at the difference function:

$$f_r(S^P, S^I) = \frac{\frac{.016 + .984(1-r)^3}{.016 + .126(1-r) + .858(1-r)^3} - \frac{.001 + .999(1-r)^3}{.001 + .027(1-r) + .243(1-r)^2 + .729(1-r)^3}}{(11)}$$

This gives the graph depicted in figure 3.

Clearly, it is not the case that $f_r(S^P, S^I) > 0$ for all values of the reliability parameter. Therefore, Bovens and Hartmann's theory has the consequence that a positively dependent set may fail to be more coherent than an independent set, even if both have the same unconditional probabilities.⁵ This seems to be a very unwelcome result.

⁵ Without the restriction that the unconditional probabilities are equal, the original example by Bovens and Hartmann (2003a: 40–43) can easily be modified so as to give a similar result as the

Fig. 3: $f_r(S^P, S^I)$.

Adding Propositions. A second general intuition about coherence is the idea that it must be possible to increase the coherence of an information set by adding propositions. Evidently, it will not always be the case that adding a proposition increases a set's coherence from an intuitive standpoint. For example, if we add a proposition that is inconsistent with the conjunction of the propositions in a set, then we should not expect coherence to increase. However, the following theorem shows that for Bovens and Hartmann's theory of coherence, adding a proposition to an information set will only increase the coherence of that set for an extremely limited number of cases (for a proof, see the appendix):

Theorem 3.1 *Adding a proposition R_{m+1} to a set $S = \{R_1, \dots, R_m\}$ that is not equivalent will increase S 's coherence iff R_{m+1} is equivalent with $\bigwedge S$, i.e., iff $p(R_1 \wedge \dots \wedge R_m | R_{m+1}) = p(R_{m+1} | R_1 \wedge \dots \wedge R_m) = 1$.*

Admittedly, adding a proposition that is equivalent with the conjunction of propositions in a set should intuitively increase the coherence of that set. But it seems that there are many other ways in which the coherence may increase. For example, consider the set S^P defined above. Now suppose that in the population under consideration all and only crows have been ringed. Furthermore, suppose that a fourth independent witness reports that the bird in question was ringed. Consider the following sets:

$$\begin{aligned} S_2^P &= \{[\text{This bird has a life-long mate}], [\text{This bird is a crow}], \\ &\quad [\text{This bird is black}], [\text{This bird is ringed}]\}; \\ S^P &= \{[\text{This bird has a life-long mate}], [\text{This bird is a crow}], \\ &\quad [\text{This bird is black}]\}. \end{aligned}$$

one above. If one sets the marginal probabilities of the propositions in their set S , which is the set consisting of independent propositions, lower than .22 or higher than .62, while leaving intact the probability model for their set S' , which is the set consisting of (some) positively dependent propositions, again no coherence ordering is possible anymore.

Alternatively, one may wish to impose different restrictions. For example, one may insist that instead of the unconditional probabilities, the prior probabilities should be equal. However, in that case similar examples are possible. For example, consider the set $S^{P'}$ with $a_0 = .001$, $a_1 = .03$ and $a_2 = 0$ and compare this with set S^I above. In this case, $S^{P'}$ and S^I have equal prior probabilities, while $S^{P'}$ is positively dependent. Still, $f_r(S^{P'}, S^I) < 0$ for $r > .9875$.

Finally, a last option is to suppose that the value of a_n must remain equal, which means, informally put, that the total area within the circles of the propositions remains equal. I cannot see any justification for that move, although in Bovens and Hartmann's example the values for a_n and a'_n are indeed equal. However, this alternative would not work either. For in this case take a set $S^{P''}$ with $a_0 = .122$, $a_1 = .149$ and $a_2 = 0$ and compare this set with S^I . It can easily be checked that $S^{P''}$ is positively dependent and that $a''_n = a'_n$. However, $f_r(S^{P''}, S^I) < 0$ for $r < .14$.

It would seem that the first set is more coherent than the second. But according to Bovens and Hartmann's account this is not the case, as figure 4 shows. Moreover, adding any number of propositions will not increase the coherence, either, as long as none of them is equivalent with the propositions already there. This, again, seems to be a very counterintuitive result.

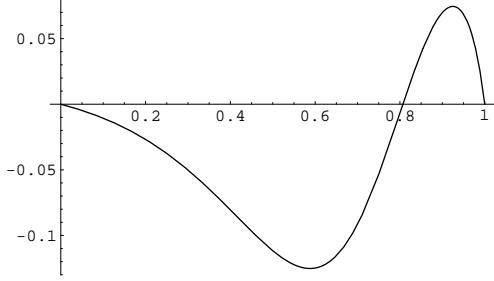


Fig. 4: $f_r(S_2^P, S^P)$.

4 Equivalent Sets and Maximal Coherence

The previous section has presented two important problems for Bovens and Hartmann's theory of coherence. Apparently in their theory it is not always the case that a positively dependent set is more coherent than an independent set, even if the propositions in both sets have the same unconditional probabilities. Moreover, adding one (or more) proposition(s) will hardly ever increase a set's coherence.

Now the question naturally arises whether or not these results are unavoidable given the model that Bovens and Hartmann have chosen for their analysis of coherence. As noted above, *Bayesian Epistemology* introduces both the idea of coherence as a confidence-boosting property and the concept of modelling coherence as a property of an information set with independent witnesses reporting on the propositions of that information set. If the possibility of examples like those above is the necessary byproduct of such an approach, then this would imply that it is a highly unsatisfying framework for analyzing coherence. This would be an important conclusion, for it would seriously challenge either the use of witnesses to analyze coherence or even the view on coherence as a confidence-boosting property. Neither is the case, as I will argue now. More specifically, I will show how Bovens and Hartmann's basic framework can serve to define at least one measure which does not share the same counterintuitive conclusions.

My basic strategy will be to deny the maximality requirement. In section 2, I have indicated that Bovens and Hartmann need an additional element in order to distinguish coherence and prior probability. As intimated there, this requirement stipulates that equivalent sets are maximally coherent *per se*, i.e., irrespective of the value of a_0 and of the number of propositions. However, the choice for this requirement seems arbitrary. Nothing in a view of coherence as a confidence-boosting property or in a model using partially reliable witnesses implies that a notion of maximal coherence should be independent of the number of propositions or the probability of their conjunction. To the contrary, there seems to be a very good reason to reject the validity of the maximality requirement, at least within this model. The reason is that this requirement implies that the posterior probability must depend on more than the coherence, the prior probability and the reliability of the witnesses.

Consider two independent witnesses with fixed reliability.⁶ Assume furthermore that they give equivalent reports, i.e., that the propositions they report as being true overlap completely. Next, a new witness enters the stage, similar to the first two in that this witness, too, has a fixed reliability and this witness, too, positively reports on a proposition that is equivalent to the other two. Since the three propositions overlap fully, $a_1 = a_2 = 0$, and it is a straightforward exercise to show that $p^*(R_1 \wedge R_2 \wedge R_3) > p^*(R_1 \wedge R_2)$ for all $a_0 \in (0, 1)$: our confidence that the information is true has gone up. This is what we would expect, of course: various independent reports that something is the case will increase our confidence that it is in fact the case. However, in this example all three of the determinants of our confidence in a set remained constant: the reliability was fixed, the prior probability remained equal, and the set was maximally coherent to begin with, so the coherence must have remained equally maximal also. But what, then, has caused our increase in confidence? Well, it seems that either we must find a fourth factor that influences the posterior probability, or relinquish the maximality requirement. If we allow that a set with three equivalent propositions is more coherent than a set with two equivalent propositions, then the problem disappears: the increase in the posterior probability is caused by an increase in the coherence. Therefore, until Bovens and Hartmann present us with this fourth factor, their account of the factors that influence our confidence in an information set is actually inconsistent with their account of coherence.⁷

Summarizing the above: not only is the maximality intuition independent of the basic framework that Bovens and Hartmann use, it is also unsatisfactory given this same basic framework. However, if it should be discarded, what other element will be able to take its place? In the next section I will not attempt to find a new requirement to replace the maximality requirement, although there may well be some quite plausible options. Instead, I will show how a relatively small adjustment in the *ceteris paribus* conditions that Bovens and Hartmann propose will make such an additional requirement superfluous. A measure of coherence can then be based directly on the revised versions of those conditions.

5 An Alternative to Bovens and Hartmann's Measure of Coherence

In section 2, I argued that no measure of coherence can be constructed on the basis of the idea that one should be more confident in a more coherent set, given equal reliability and equal prior probability. The reason for this is that prior probability and coherence are not independent notions. Of course this need not be an impediment to the construction of a measure of coherence: as intimated above, Bovens and Hartmann simply add a requirement (to wit, the maximality requirement) from which their measure follows quite straightforwardly. However, the last section has criticized this intuition. So it seems that at least one way to proceed would be to search for a more plausible intuition and construct a measure on that basis. Alternatively, one may try to find three determinants of confidence that *can* be modelled as independent of one another. This is the approach that I will pursue in the rest of this paper.

Before I start, let me be clear about the goal of this project. Emphatically, I

⁶ One may assume that the witnesses are actual persons reporting on propositions and that long experience with these people has taught us the true- and false-positive rates of their reports. Alternatively, one may assume the witnesses to be independent measurement instruments of which we likewise know the number of true and false positives.

⁷ The maximality requirement faces other, more general, problems also. For an argument against the maximality requirement that does not depend on Bovens and Hartmann's theory, see Douven and Meijs (2004).

am not trying to present a foundation for a new measure of coherence. Instead, I am merely trying to show how Bovens and Hartmann's difference function can be adjusted in order to evade the problems that I have raised in the previous sections. In doing so I will try to deviate as little as possible from Bovens and Hartmann's original model. Those who did not find that model or, for that matter, the notion of coherence as a confidence-boosting property, plausible will not find an answer to their misgivings in this paper. Instead, it addresses those who were equally attracted by Bovens and Hartmann's original theory as worried by the examples put forward against that theory in this and earlier papers.

In this revised version of Bovens and Hartmann's theory, coherence will still be a property of an information set and it will remain a confidence-boosting property in the sense that the more coherent the story is, the more confident we will be that the information is true, *ceteris paribus*. What will be different, however, is the explication of the *ceteris paribus* condition. Instead of requiring equal *prior* probability, I will require equal *unconditional* probabilities. That is, according to me, one should be more confident in an information set if it is more coherent, given equal reliability and equal *unconditional* probabilities. This will make our confidence in an information set depend on the unconditional probabilities, the witness reliability and the coherence, but no longer on the prior probability. In that case, a set becomes more coherent if its unconditional probabilities and the reliability of the witnesses remain equal while the value of $p^*(R_1 \wedge \dots \wedge R_n)$ increases. And similarly, if the unconditional probabilities of an independent set increase without altering the fact that they are independent, we will become more convinced that the information is true, while the coherence will be left untouched.

But, rather surprisingly perhaps, it has now become possible to construct a measure on the basis of only this description of coherence as a confidence-boosting property. For while it is certainly not the case that prior probability and coherence are independent, it is not evident at all that unconditional probabilities and coherence cannot be modelled as independent.

Above I have argued that sets consisting of independent propositions have neutral coherence, no matter how plausible they are. This at least seems to strongly suggest that the unconditional probabilities do not determine the degree of a set's coherence. Now if one agrees with this, then it becomes possible to measure coherence by comparing the actual probability boost with the probability boost the set would have gotten had all its propositions been independent (but with the same unconditional probabilities and the same witness reliability).⁸

To formalize this idea, again consider the set \mathbf{S} of information sets. But now define the notion $p_{\text{ind}}^*(R_1 \wedge \dots \wedge R_n)$ as the posterior probability of the information set $S = \{R_1, \dots, R_n\}$, had the information come in as independent propositions with the same unconditional probabilities. Call the set consisting of the same propositions as S , with the same unconditional probabilities, but in which these propositions are independent, set S^i and let $b_{\text{ind}}(R_1 \wedge \dots \wedge R_n)$ be the probability boost of the independent set. Next, let a_j^i stand for the sum of the probabilities of the conjunctions consisting of $n - j$ elements of S^i and the negations of all the remaining elements of S^i . From these elements, one can construct a new coherence

⁸ One may feel that the proposal to compare a set's coherence with the coherence it would have gotten had the information come in in an independent fashion, does not immediately follow from the above *ceteris paribus* conditions. Although I would say that it seems to be the only possible option, the worry can be parried in another way as well. Just as Bovens and Hartmann add the maximality requirement, one may add, in this case, what may be called the *independence requirement*, which says that independent sets should have neutral coherence. From this requirement the measure proposed below follows smoothly.

measure:

$$c_r^i(S) =_{df} \frac{b(R_1 \wedge \dots \wedge R_n)}{b_{\text{ind}}(R_1 \wedge \dots \wedge R_n)} = \frac{a_0}{a_0^i} \frac{\sum_{i=0}^n a_i^i (1-r)^i}{\sum_{i=0}^n a_i (1-r)^i}, \quad (12)$$

and a new difference function:

$$f_r^i(S, S') = c_r^i(S) - c_r^i(S'). \quad (13)$$

As in the original case, $S \succcurlyeq_i S'$ iff $f_r^i(S, S') \geq 0$ for all values of $r \in (0, 1)$.

At first glance, this account has some important advantages over the old one. For one, it now becomes possible to define the properties of coherence and incoherence of a set, which I shall call i-coherence:

Definition 5.1 A set S is i-coherent iff $c_r^i(S) > 0$ for all values of $r \in (0, 1)$; it is i-incoherent iff $c_r^i(S) < 0$ for all values of $r \in (0, 1)$.

Also, let me stress that this alternative measure leaves intact most of Bovens and Hartmann's model. Firstly, coherence is still a confidence-boosting property of an information set. Secondly, the independence of the propositions of the information set is still analyzed in terms of independent witnesses reporting on propositions in an information set. And finally, the coherence relation ' \succcurlyeq_i ' remains a quasi-ordering.

But more importantly, the new measure gives the intuitively correct answers to the examples given in section 3. Figures (5) and (6) show the graphs of $f_r^i(S^P, S^I)$ and $f_r^i(S_2^P, S^P)$. Clearly, both graphs are positive for all values of the reliability parameter. Therefore, this revised version of Bovens and Hartmann's measure of coherence agrees with our intuitive judgements that the positively dependent set S^P defined above is more coherent than the independent set S^I and that S_2^P is more coherent than S^P . Moreover, this measure gives the intuitively correct answers to the examples brought forward against the original difference function in Douven and Meijs (2004) and Meijs and Douven (2004). And finally, it agrees with almost all the examples that Bovens and Hartmann discuss in their (2003a) and (2003b).

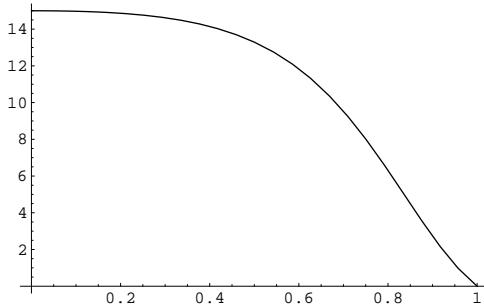
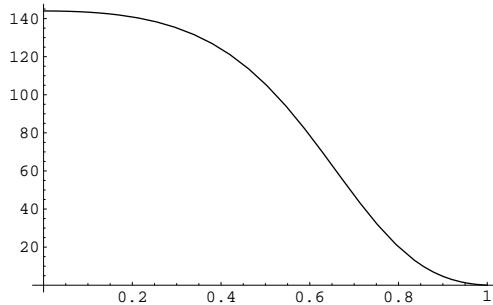


Fig. 5: $f_r^i(S^P, S^I)$.

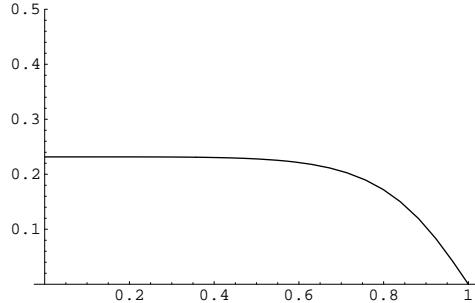
To be precise, the new measure disagrees with Bovens and Hartmann's intuitions in only two examples. The first of these is the example that Bovens and Hartmann (2003a: 50–53) use to criticize Fitelson's (2003) measure. I will not discuss this example any further, since, contrary to Bovens and Hartmann's discussion, I do not believe that it is such an intuitively clear example. For a criticism of this example, see Douven and Meijs (2004).

The second example on which the two measures differ might be considered more important, since it is the only example that Bovens and Hartmann give of two sets between which no coherence ordering is possible. Here is the example. A murder has been committed in Tokyo but the corpse is still to be found. A 100-square grid

Fig. 6: $f_r^i(S_2^P, S^P)$.

has been drawn over the map of the city and initially the corpse is equally likely to be found in any of the squares. Now consider two different situations, δ and ϵ , in each of which two independent and equally, but partially, reliable witnesses report about the location of the corpse. In situation δ the two witnesses report squares 41–60 and 51–70. In situation ϵ the witnesses report squares 39–61 and 50–72.

Define ' S^δ ' and ' S^ϵ ' as the information sets containing the witness reports of situations δ and ϵ , respectively. Given the above probabilities, we have $a_0^\delta = .10$, $a_1^\delta = .20$, $a_0^\epsilon = .12$, and $a_1^\epsilon = .22$. Since $f(S^\delta, S^\epsilon)$ is positive for values of r between 0 and, approximately, .7 and negative for values of r between, approximately, .7 and 1, we have neither $S^\delta \succcurlyeq S^\epsilon$ nor $S^\epsilon \succcurlyeq S^\delta$. However, as figure 7 shows, $f_r^i(S^\delta, S^\epsilon)$ is larger than 0 for all values of r , and therefore, $S^\delta \succcurlyeq_i S^\epsilon$.

Fig. 7: $f_r^i(S^\delta, S^\epsilon)$.

According to the difference function there is no fact of the matter as to which set is more coherent, and Bovens and Hartmann believe that this constitutes a point in favor of their theory. In such cases, they think, a theory of coherence should withhold judgement. But it seems telling that almost all the other recently proposed probabilistic measures of coherence side with $f_r^i(S^\delta, S^\epsilon)$. Fitelson's, Shogenji's and the three measures proposed in Douven and Meijs (2004) all point to S^δ as being more coherent than S^ϵ .⁹ Only according to Olsson's (2002) tentative proposal is the opposite the case.¹⁰ So there is at least some reason to argue that S^δ is more coherent than S^ϵ . Moreover, this conclusion is not incompatible with Bovens and Hartmann's claim that in this case it is intuitively not clear which set is more

⁹ Shogenji's (1999) measure yields $S^\delta = 2.5$ and $S^\epsilon \approx 2.27$; Fitelson's (2003) measure yields $S^\delta = .6$ and $S^\epsilon \approx .57$, the three measures discussed in Douven and Meijs (2004) yield $\mathcal{C}_d(S^\delta) = .3$, $\mathcal{C}_d(S^\epsilon) \approx .29$, $\mathcal{C}_r(S^\delta) = 2.5$, $\mathcal{C}_r(S^\epsilon) \approx 2.27$, $\mathcal{C}_l(S^\delta) = 4.0$, $\mathcal{C}_l(S^\epsilon) \approx 3.6$.

¹⁰ This measure yields $S^\delta \approx .333$ and $S^\epsilon \approx .353$.

coherent. We may intuitively be unable to see the difference in coherence for the simple reason that it is too small for us to recognize.

As can be seen from the values of the respective measures, according to all of the measures mentioned above the differences in coherence between the two sets are indeed minute, which would explain our being unable to point to the most coherent set. Of course, Bovens and Hartmann's framework does not supply a mechanism with which to decide *how much* more coherent a set is than another, which makes it impossible to distinguish between intuitively very clear cases and intuitively less clear cases. This is just a natural result of their approach. However, it is quite easy to add such an element to $f_r^i(S, S')$. For example, one could take the maximum value of $f_r^i(S, S')$ or the area beneath its graph to be indicative of the degree to which S is more coherent than S' (assuming $f_r^i(S, S') \geq 0$ for all r). In the case of S^δ and S^ϵ , both proposals would lead to the conclusion that the difference in coherence between S^δ and S^ϵ is much smaller than in any of the other examples.

As a last remark, let me note that although this measure does not satisfy the maximality requirement, it does satisfy a less stringent condition, to wit, that sets consisting of equivalent propositions are maximally coherent, given constant unconditional probabilities and equal reliability of the witnesses (for a proof, see the appendix):

Theorem 5.1 *Any equivalent set $S = \{R_1, \dots, R_n\}$ with unconditional probabilities $p(R_i) = \alpha$ is more i-coherent than any nonequivalent set S' with the same number of propositions and the same unconditional probabilities.*

Thus, given two sets with the same unconditional probabilities for all the propositions, if only one of them is equivalent, it will be more i-coherent than the other. Or, alternatively, given fixed unconditional probabilities, the most i-coherent way information can accumulate is if the propositions are all equivalent.

6 Conclusion

In this paper I have shown that Bovens and Hartmann's theory of coherence runs into some serious problems. It turned out that in some cases a positively dependent set is not more coherent than an independent set, even if both consist of the same number of propositions with the same unconditional probabilities. Also, it turned out that adding arbitrarily many propositions to a set cannot make it more coherent if these propositions are not equivalent with the conjunction of the propositions already there.

One of the suggestions of this paper has been that the maximality requirement should be discarded. Not only does it lead to a measure that is intuitively unsatisfying, it is also at odds with Bovens and Hartmann's view on coherence as a confidence-boosting property. As I have remarked in section 4, the maximality requirement leads to the conclusion that no matter how many equivalent propositions a set consists of, it will be equally coherent to any other set with equivalent propositions. In light of the view on coherence as a confidence-boosting property, this does not make sense.

But the main conclusion of this paper must be that a view on coherence as a confidence-boosting property does not directly determine a measure of coherence, even if one accepts the independent-witness model. In section 5 I have discussed one alternative, and it seems safe to conclude that it behaves in a much more satisfactory way than Bovens and Hartmann's proposal.

Acknowledgements

I am greatly indebted to Luc Bovens, Stephan Hartmann and Igor Douven for their many helpful comments on earlier versions of this paper.

References

- Akiba, K. [2000] “Shogenji’s Probabilistic Measure of Coherence is Incoherent,” *Analysis* 60: 356–359.
- BonJour, L. [1985] *The Structure of Empirical Knowledge*, Cambridge MA: Harvard University Press.
- Bovens, L. and S. Hartmann [2003a] *Bayesian Epistemology*, Oxford: Oxford University Press.
- Bovens, L. and S. Hartmann [2003b] “Solving the Riddle of Coherence,” *Mind* 112: 601–633.
- Douven, I. and W. Meijs [2004] “Measuring Coherence,” *Synthese*, in press.
- Eells, E. and B. Fitelson [2002] : “Symmetries and Asymmetries in Evidential Support”, *Philosophical Studies*, 107: 129–142.
- Fitelson, B. [2003] “A Probabilistic Theory of Coherence,” *Analysis* 63: 194–199.
- Meijs, W. and I. Douven [2004] “Bovens and Hartmann on Coherence,” *Mind*, in press.
- Olsson, E. [2002] “What is the Problem of Coherence and Truth?”, *Journal of Philosophy* 94: 246–272.
- Shogenji, T. [1999] “Is Coherence Truth-Conducive?”, *Analysis* 59: 338–345.

Appendix: Proofs of Propositions 3.1 and 5.1

Theorem 3.1 Adding a proposition R_{m+1} to a set $S = \{R_1, \dots, R_m\}$ that is not equivalent will increase S 's coherence iff R_{m+1} is equivalent with $\bigwedge S$, i.e., iff $p(R_1 \wedge \dots \wedge R_m \mid R_{m+1}) = p(R_{m+1} \mid R_1 \wedge \dots \wedge R_m) = 1$.

Proof: I will start by showing that adding a proposition R_{m+1} to any set $S = \{R_1, \dots, R_m\}$ will not increase S 's coherence if R_{m+1} is not equivalent with S . To this end, consider two sets $S = \{R_1, \dots, R_m\}$ and $S' = \{R_1, \dots, R_m, R_n\}$, where $n = m + 1$. For convenience, substitute b_i 's for the a'_i 's of S' .

Either $b_0 = 0$ or $b_0 > 0$. In the first case, $f(S', S)$ is not defined, and, therefore, S' cannot be more coherent than S . Thus, b_0 must be larger than 0.

For it to be the case that S' is not more coherent than S , there must be at least one value r_0 such that $f_{r_0}(S', S) < 0$. Furthermore, note that for this to be the case it would suffice if the derivative of $f_r(S', S)$ is negative in $r = 0$, since this implies that $f_r(S', S) < 0$ for values of r approximating 0.¹¹ To find the expression for the derivative of $f_r(S', S)$ for $r = 0$, first consider $c_r(S^*)$ for a set $S^* = \{R_1, \dots, R_K\}$ with K propositions:

$$c(S^*) = \frac{a_0 + (1 - a_0)(1 - r)^K}{\sum_{i=0}^K a_i (1 - r)^i},$$

where I have dropped the subscript r . Next, substitute $1 - r = x$. Then:

$$\frac{d c(S^*)}{dr} = -\frac{d c(x)}{dx}.$$

For r approximating 0, x approximates 1. Next, calculate $(\frac{d c(x)}{dx})_{x=1}$ using $\sum_{i=0}^K a_i = 1$:

$$\begin{aligned} \left(\frac{d c(x)}{dx} \right)_{x=1} &= \left(\left(\frac{1}{\sum_{i=0}^K a_i x^i} \right) K(1 - a_0)x^{K-1} \right)_{x=1} - \\ &\quad \left((a_0 + (1 - a_0)x^K) \left(\frac{\sum_{i=0}^K a_i i x^{i-1}}{\left(\sum_{i=0}^K a_i x^i \right)^2} \right) \right)_{x=1} \\ &= \left(\frac{1}{\sum_{i=0}^K a_i} \right) K(1 - a_0) - (a_0 + (1 - a_0)) \left(\frac{\sum_{i=0}^K a_i i}{\left(\sum_{i=0}^K a_i \right)^2} \right) \\ &= K(1 - a_0) - \left(\sum_{i=0}^K a_i i \right) \\ &= K(1 - a_0) - \left(\sum_{i=0}^{K-1} i a_i + K a_K \right) \\ &= K(1 - a_0) - \left(\sum_{i=0}^{K-1} i a_i + K \left(1 - \sum_{i=0}^{K-1} a_i \right) \right) \\ &= (K - K a_0) - \left(K + \sum_{i=0}^{K-1} (i - K) a_i \right) \end{aligned}$$

¹¹ I thank Stephan Hartmann for pointing out that this approach is much easier than the one I originally had in mind.

$$\begin{aligned}
&= -Ka_0 - \left(\sum_{i=0}^{K-1} (i-K)a_i \right) \\
&= - \sum_{i=1}^{K-1} (i-K)a_i.
\end{aligned}$$

Since $\frac{dc(S^*)}{dr} = -\frac{dc(x)}{dx}$, it follows that

$$\left(\frac{dc(S^*)}{dr} \right)_{r=0} = \sum_{i=1}^{K-1} (i-K)a_i. \quad (1)$$

Using equation (1), the derivative of $f(S', S)$ for sets $S = \{R_1, \dots, R_m\}$ and $S' = \{R_1, \dots, R_m, R_n\}$ equals

$$\begin{aligned}
\left(\frac{df(S', S)}{dr} \right)_{r=0} &= \sum_{i=1}^{n-1} (i-n)b_i - \sum_{i=1}^{m-1} (i-m)a_i \\
&= \sum_{i=1}^{m-1} (m-i)a_i - \sum_{i=1}^{n-1} (n-i)b_i.
\end{aligned} \quad (2)$$

In order to compare the two sums above, it is necessary to derive a relation between the a_i 's and the b_i 's. Start with dividing the b_i 's in two parts: $b_i^{R_n}$ is the part of b_i where R_n is true and $b_i^{\neg R_n}$ is the part of b_i where R_n is false. In that case, $b_i^{R_n} + b_i^{\neg R_n} = b_i$ and $b_0^{\neg R_n} = b_n^{R_n} = 0$. Also:

$$\sum_{i=0}^n b_i^{\neg R_n} = \sum_{i=1}^n b_i^{\neg R_n} = 1 - p(R_n). \quad (3)$$

Furthermore, b_0 can be expressed in terms of a_0 and $b_1^{R_n}$ by using the law of total probability:

$$\begin{aligned}
b_0 &= p(R_1 \wedge \dots \wedge R_m \wedge R_n) \\
&= p(R_1 \wedge \dots \wedge R_m) - p(R_1 \wedge \dots \wedge R_m \wedge \neg R_n) \\
&= a_0 - b_1^{R_n}.
\end{aligned} \quad (4)$$

The same type of relation can be derived for all b_i . By definition, $b_i^{R_n}$ is the sum of all possible combinations in which R_n is true and precisely i propositions are false:

$$b_i^{R_n} = \sum_{\substack{i\text{-many } \neg R_k\text{'s}, \\ (m-i)\text{-many } R_k\text{'s}}} p(R_1 \wedge \dots \wedge R_m \wedge R_n)$$

with $k \leq m$. Again by the law of total probability,

$$\begin{aligned}
b_i^{R_n} &= \sum_{\substack{i\text{-many } \neg R_k\text{'s}, \\ (m-i)\text{-many } R_k\text{'s}}} (p(R_1 \wedge \dots \wedge R_m) - p(R_1 \wedge \dots \wedge R_m \wedge \neg R_n)) \\
&= \sum_{\substack{i\text{-many } \neg R_k\text{'s}, \\ (m-i)\text{-many } R_k\text{'s}}} p(R_1 \wedge \dots \wedge R_m) - \sum_{\substack{i\text{-many } \neg R_k\text{'s}, \\ (m-i)\text{-many } R_k\text{'s}}} p(R_1 \wedge \dots \wedge R_m \wedge \neg R_n).
\end{aligned}$$

The first term on the right-hand side of the last equation equals a_i : it is the sum of all the joint probabilities of all combinations of set S where i propositions are false and all others are true. The second term equals $b_{i+1}^{\neg R_n}$: it is the sum of the

joint probabilities of all combinations of set S' in which R_n and precisely i other propositions are false. Therefore:

$$b_i^{R_n} = a_i - b_{i+1}^{\neg R_n} \quad (5)$$

and

$$b_i = b_i^{R_n} + b_i^{\neg R_n} = a_i - b_{i+1}^{\neg R_n} + b_i^{\neg R_n}. \quad (6)$$

Substituting this in expression (2) and using $n = m + 1$ gives:

$$\begin{aligned} & \sum_{i=1}^{m-1} (m-i)a_i - \sum_{i=1}^{n-1} (n-i)b_i \\ = & \sum_{i=1}^{m-1} (m-i)a_i - \sum_{i=1}^{n-1} (n-i) (a_i + b_i^{\neg R_n} - b_{i+1}^{\neg R_n}) \\ = & \sum_{i=1}^{m-1} (m-i)a_i - \sum_{i=1}^m (m+1-i)a_i - \sum_{i=1}^m (n-i)b_i^{\neg R_n} + \sum_{i=1}^m (n-i)b_{i+1}^{\neg R_n} \\ = & \sum_{i=1}^{m-1} ((m-i)a_i - (m+1-i)a_i) - a_m - \sum_{i=1}^m (n-i)b_i^{\neg R_n} + \sum_{i=2}^{m+1} (n-i+1)b_i^{\neg R_n} \\ = & \sum_{i=1}^{m-1} (-a_i) - a_m - \sum_{i=1}^m (n-i)b_i^{\neg R_n} + \sum_{i=2}^{m+1} (n-i+1)b_i^{\neg R_n} \\ = & \sum_{i=1}^m (-a_i) - (n-1)b_1^{\neg R_n} - \sum_{i=2}^m (n-i)b_i^{\neg R_n} + \sum_{i=2}^m (n-i+1)b_i^{\neg R_n} + b_n^{\neg R_n} \\ = & -\sum_{i=1}^m a_i - mb_1^{\neg R_n} + \sum_{i=2}^m b_i^{\neg R_n} + b_n^{\neg R_n} \\ = & -\sum_{i=1}^m a_i + \sum_{i=1}^n b_i^{\neg R_n} - b_1^{\neg R_n} - mb_1^{\neg R_n}. \end{aligned}$$

Filling in equations (3) and (4) and substituting $\sum_{i=1}^m a_i = 1 - a_0$, it is possible to derive a simple expression for the derivative of $f(S', S)$ for $r = 0$:

$$\begin{aligned} \left(\frac{d f(S', S)}{dr} \right)_{r=0} &= -1 + b_0 + b_1^{\neg R_n} + 1 - p(R_n) - b_1^{\neg R_n} - mb_1^{\neg R_n} \\ &= b_0 - p(R_n) - mb_1^{\neg R_n} \\ &= p(R_1 \wedge \dots \wedge R_m \wedge R_n) - p(R_n) - \\ &\quad mp(R_1 \wedge \dots \wedge R_m \wedge \neg R_n) \\ &= p(R_n)(p(R_1 \wedge \dots \wedge R_m | R_n) - 1) - \\ &\quad mp(\neg R_n)p(R_1 \wedge \dots \wedge R_m | \neg R_n). \end{aligned}$$

From this it follows directly that the derivative of $f(S', S)$ for $r = 0$ is negative if $p(R_1 \wedge \dots \wedge R_m | R_n) < 1$, or if $p(R_1 \wedge \dots \wedge R_m | \neg R_n) > 0$. Therefore, adding a proposition R_{m+1} to a set of propositions $S = \{R_1, \dots, R_m\}$ will not increase its coherence if R_{m+1} is not equivalent with S .

If R_{m+1} is equivalent with S , coherence will increase if set S is nonequivalent. For consider the coherence function for set $S' = \{R_1, \dots, R_m, R_n\}$:

$$c_r(S') = \frac{b_0 + (1-b_0)(1-r)^n}{\sum_{i=0}^n b_i(1-r)^i}.$$

If R_n is equivalent with $R_1 \wedge \dots \wedge R_m$, then $b_0 = a_0$, $b_1 = 0$ and $b_{i+1} = a_i$ for $i \geq 1$. Therefore:

$$\begin{aligned} c_r(S') &= \frac{a_0 + (1 - a_0)(1 - r)^n}{a_0 + \sum_{i=2}^n a_{i-1}(1 - r)^i} \\ &= \frac{a_0 + (1 - a_0)(1 - r)^n}{a_0 + \sum_{i=1}^m a_i(1 - r)^{i+1}} \\ &= \frac{\frac{a_0}{(1-r)} + (1 - a_0)(1 - r)^m}{\frac{a_0}{(1-r)} + \sum_{i=1}^m a_i(1 - r)^i} \end{aligned}$$

It is fairly easy to show that for all $\alpha, \beta, \gamma, \delta \in (0, 1)$, if $\gamma < \delta$,

$$\frac{\frac{\alpha}{\beta} + \gamma}{\frac{\alpha}{\beta} + \delta} > \frac{\alpha + \gamma}{\alpha + \delta}.$$

If S is nonequivalent, then

$$\sum_{i=1}^m a_i(1 - r)^m < \sum_{i=1}^m a_i(1 - r)^i,$$

and thus $(1 - a_0)(1 - r)^m < \sum_{i=1}^m a_i(1 - r)^i$. Therefore, for all $0 < r < 1$ and all $0 < a_0 < 1$, we have

$$\begin{aligned} c_r(S') &> \frac{a_0 + (1 - a_0)(1 - r)^m}{a_0 + \sum_{i=1}^m a_i(1 - r)^i} \\ &= c_r(S). \end{aligned}$$

Therefore, $f_r(S', S) =_{df} c_r(S') - c_r(S) > 0$ for all values of r and thus S' is more coherent than S . ■

Theorem 5.1 *Any equivalent set $S = \{R_1, \dots, R_n\}$ with unconditional probabilities $p(R_i) = \alpha$ is more i-coherent than any nonequivalent set S' with the same number of propositions and the same unconditional probabilities.*

Proof: Consider a set $S = \{R_1, \dots, R_n\}$ consisting only of equivalent propositions with probability $p(R_i) = \alpha$. Next consider a second set $S' = \{R'_1, \dots, R'_n\}$ with the same number of propositions, each of which has unconditional probability $p(R'_i) = \alpha$, but in which at least one proposition or subset of propositions is independent of all the others.

First, note that since all propositions in both sets have the same unconditional probabilities, the posterior probabilities of the independent sets are equal: $p^*(S^I) = p^*(S'^I)$.

Next, note that in this case $a_0 > a'_0$. Therefore, the actual posterior probability of the first set is larger than that of the second: $p^*(S) > p^*(S')$. Thus, $c_r^i(S) > c_r^i(S')$ for all r , and therefore $f_r^i(S, S')$ is positive for all values of the reliability parameter r . ■