

# New Hope for Shogenji's Coherence Measure\*

Jonah N. Schupbach<sup>†</sup>

October 14, 2009

[forthcoming in *BJPS*]

## Abstract

I show that the two most devastating objections to Shogenji's formal account of coherence necessarily involve information sets of cardinality  $n \geq 3$ . Given this, I surmise that the problem with Shogenji's measure has more to do with his means of *generalizing* the measure than with the measure itself. I defend this claim by offering an alternative generalization of Shogenji's measure. This alternative retains the intuitive merits of the original measure while avoiding both of the relevant problems that befall it. In the light of all of this, I suggest that there is new hope for Shogenji's analysis; Shogenji's early and influential attempt at measuring coherence, when generalized in a *subset-sensitive* way, is able to clear its most troubling objections.

---

\*I owe special thanks to Luc Bovens, Stephan Hartmann, and Jan Sprenger for helpful conversations pertaining to this paper. I also am grateful to Tilburg University's Center for Logic and Philosophy of Science (TiLPS) for supporting me with a research fellowship during the time that I wrote this paper.

<sup>†</sup>University of Pittsburgh, Department of History & Philosophy of Science; 1017 Cathedral of Learning; Pittsburgh, Pennsylvania, USA; 15260.

In a paper written over ten years ago now, Tomoji Shogenji (1999) initiated a flurry of work on formal analyses and measures of the epistemological concept of coherence. Countless articles have now been written in response to Shogenji's paper (for a small but important sample, see Akiba, 2000; Fitelson, 2003; Glass, 2005; and Olsson, 2002), books and dissertations have appeared on the topic (Bovens and Hartmann, 2003; Olsson, 2005; Meijs, 2005), impossibility results have been proposed for Bayesian accounts of coherence such as Shogenji's (Bovens and Hartmann, 2003; Olsson, 2005) and rebuttals to these have been attempted (Douven and Meijs, 2007b; Schupbach, 2008; Wheeler, 2009). In short, Shogenji's paper has sparked a fascinating and flourishing new sub-discipline within formal epistemology.

The fact that Shogenji's article has had a high impact is indubitable. More controversial is the value of the positive formal account that he presents in that article. At the very least, Shogenji's analysis of coherence still has the merits of being one of the most uncomplicated, accessible accounts on the market. The basic idea is very simple: the coherence of an information set is represented by the positive relevance among the members of that set. For any set of cardinality two  $S = \{R_1, R_2\}$ , coherence amounts to the following:

$$Coh(S) = \frac{Pr(R_1 \wedge R_2)}{Pr(R_1) \times Pr(R_2)}$$

Additionally, Shogenji offers the following straightforward generalization of this measure to sets of cardinality greater than or equal to two  $S = \{R_1, R_2, \dots, R_n\}$ :

$$Coh(S) = \frac{Pr(R_1 \wedge R_2 \wedge \dots \wedge R_n)}{Pr(R_1) \times Pr(R_2) \times \dots \times Pr(R_n)}$$

The function *Coh* compares the joint probability of the members of  $S$  (the numerator) with the value that this joint probability would take were these members statistically independent of one another (the denominator). It thereby provides a measure of the degree to which the members of this information set are statistically dependent on, or relevant to, one another.

In spite of - and in part, *because of* - its simplicity, *Coh* has been the target of numerous critiques. The prevailing opinion today is that *Coh* ultimately fails to capture formally the concept of coherence. Accordingly, a host of alternative coherence measures have been proposed by Bayesians. Such measures have grown increasingly complex beginning with Olsson's set-theoretic notion of coherence as relative overlap (Olsson, 2002), through Fitelson's measure of coherence as a function of Kemeny-Oppenheim degrees of factual support (Fitelson, 2003, 2004), and up to Douven and Meijs' general schema for a measure of coherence as a weighted average of the degrees of mutual support between all of the non-empty, mutually exclusive subsets of  $S$  (Douven and Meijs, 2007a). As such formal accounts become more abundant and complicated, it becomes increasingly difficult to judge their relative, intuitive merits. As a result, the pursuit of an accurate formal account of coherence seems to have come to a standstill. It should come as a relief then that this paper argues that Bayesians working on coherence should return to

their simple beginnings. I claim that Shogenji's transparent and accessible notion of coherence as positive relevance is still very much a viable option insofar as the objections raised against it can be straightforwardly avoided.

I proceed in Section 1 by briefly reviewing the objections to Shogenji's measure. I have my own devices for providing such a review: I will ultimately aim to show that those critiques that remain worrisome for the measure are only applicable to its *generalized* version - i.e., these problems necessarily involve sets of cardinality  $n > 2$ . Given this conclusion, I will introduce in Section 2 an alternative generalization of Shogenji's measure that retains the intuitive merits of the original measure while avoiding its downfalls. Finally, in Section 3, I will conclude that Shogenji's measure (in this alternatively generalized form) is very much still a viable option as a formal analysis of coherence, and I will discuss future lines of research pertaining to this measure.

## 1 Challenges to *Coh*

Surprisingly, in digging through the extensive literature on formal accounts of coherence, one finds that many of the well-known objections to Shogenji's formal account have been met with convincing rebuttals (from Shogenji as well as others). To take two such examples: **(1)** Akiba (2000) objects to Shogenji's measure on the grounds that it does not assign equal coherence to sets that have the same logical consequences and given that it doesn't apply to single beliefs (i.e., information sets of cardinality one). But Shogenji (2001) and Olsson (2005, pp. 101-102) both offer convincing rebuttals to these points by arguing that it would not be desirable for Shogenji's measure of coherence to have either of these features.

The epistemological concept of coherence for which Shogenji is trying to account is one that intuitively is distinct from any potential notion of coherence that pertains to singletons; it is instead a *relational* notion having to do with the interconnectedness of *multiple* propositions. Additionally, according to this epistemic notion, it is manifestly not true that sets with the same logical consequences ought to take the same degree of coherence. Shogenji (2001, p. 148) shows this via the following example. Consider the following three statements:

- A The fossil was deposited 64-to-66 million years ago
- B The fossil was deposited 63-to-67 million years ago
- C The fossil was deposited  $> 10$  million years ago

A logically entails both B and C; thus, the sets {A,B} and {A,C} have the same logical consequences. Nonetheless, there is a clear sense in which {A,B} is much more coherent than {A,C}.

**(2)** Fitelson (2003, pp. 196-197) objects to Shogenji's measure because it fails to satisfy Fitelson's own "logical equivalence / maximality" desideratum for any measure of coherence.<sup>1</sup> Fitelson's desideratum holds that any adequate measure

---

<sup>1</sup>Bovens and Hartmann (2003, p. 50) put forward a putative counterexample to Shogenji's account that ultimately turns on the same point as Fitelson's objection.

of coherence will have a maximal value and that it will take this value precisely whenever the information set in question is made up entirely of logically equivalent statements. However, *Coh* does not take a maximal value in such a case; instead, it gives the following result (where  $Pr(R_i) = p$ , for all  $i$ ):

$$Coh(S) = \frac{Pr(R_1 \wedge R_2 \wedge \dots \wedge R_n)}{Pr(R_1) \times Pr(R_2) \times \dots \times Pr(R_n)} = \frac{p}{p^n} = \frac{1}{p^{n-1}}$$

Responding to this challenge, Meijs argues that Fitelson’s desideratum is intuitively incompatible with any formal measure that takes coherence to be a function of positive relevance.<sup>2</sup> That is, one cannot satisfy Fitelson’s desideratum with such a measure of coherence without running into deeply counterintuitive results. Furthermore, as Douven and Meijs (2007a, p. 421) point out, it is not so clear why one would want a coherence measure to satisfy Fitelson’s desideratum anyway. If one takes coherence to be a matter of positive relevance among the members of an information set, then one’s coherence measure should provide a maximal value *relative to a particular set* just when that set’s members are maximally relevant to each other. But a set’s degree of relevance remains a function of the joint probability that it would have were it made up of entirely independent bits, and thus it can vary between sets (even when those sets are made up entirely of logically equivalent propositions). Consequently, the degree of coherence should also be allowed to vary between such sets.

Despite *Coh*’s success in meeting these challenges, there remain at least two important problems that this measure is not able to meet: the *Depth Problem* and the *Problem of Irrelevant Addition*. The first of these is, by now, a common objection to *Coh* while, to my knowledge, the latter problem is original to this paper.

## 1.1 The Depth Problem

Fitelson (2003, pp. 196-197) first raises this important objection to Shogenji’s analysis, exploiting this measure’s relative lack of probabilistic depth. As he writes (p. 197):

Shogenji’s measure is based only on the  $n$ -wise (in)dependence of the set  $E$ . It is well known that a set  $E$  can be  $j$ -wise independent, but not  $i$ -wise independent, for any  $i \neq j$ ... Since Shogenji’s measure is based only on  $n$ -wise independence (dependence), in cases where a set is  $n$ -wise independent (dependent), but not  $j$ -wise (for some  $j \neq n$ ) independent (dependent), Shogenji’s measure does not take into account the ‘mixed’ nature of the coherence (incoherence) of  $E$  (and its subsets), and it judges  $E$  as having the same degree of coherence (incoherence) as a fully independent (or fully dependent) set.

To put this objection more simply, Shogenji’s measure does not dig deeply enough into the probabilistic information of the scenario. This worrisome feature turns

<sup>2</sup>See Meijs (2005, pp. 58-61) and Meijs (2006, pp. 237-240).

into a fault in light of the fact that it leads *Coh* to giving counterintuitive results. Consider the following case:<sup>3</sup>

Police investigators have collected together eight suspects for a robbery. The robber is certainly one of these eight suspects, and each of the suspects are equally likely to have committed the crime. Additionally, there are three independent witnesses to the robbery. The witnesses are called in to the police station to view the lineup of suspects and to pick out the criminal. In scenario one, after viewing the lineup, the three witnesses offer the following information  $\mathbf{E} = \{W_1, W_2, W_3\}$ :

$W_1$ : The criminal was either suspect 1, 2, or 3.

$W_2$ : The criminal was either suspect 1, 2, or 4.

$W_3$ : The criminal was either suspect 1, 3, or 4.

In scenario two, after viewing the lineup, the three witnesses offer the following information  $\mathbf{E}' = \{W'_1, W'_2, W'_3\}$ :

$W'_1$ : The criminal was either suspect 1, 2, or 3.

$W'_2$ : The criminal was either suspect 1, 4, or 5.

$W'_3$ : The criminal was either suspect 1, 6, or 7.

Intuitively, the testimonies of the witnesses in the first scenario “hang together” much more so than the testimonies of the witnesses in the second scenario; i.e.,  $\mathbf{E}$  is more coherent than  $\mathbf{E}'$ . Yet, Shogenji’s coherence measure ranks them equally coherent:

$$Coh_S(\mathbf{E}) = \frac{Pr(W_1 \wedge W_2 \wedge W_3)}{Pr(W_1) \times Pr(W_2) \times Pr(W_3)} = \frac{1/8}{(3/8)^3} = 64/27$$

$$Coh_S(\mathbf{E}') = \frac{Pr(W'_1 \wedge W'_2 \wedge W'_3)}{Pr(W'_1) \times Pr(W'_2) \times Pr(W'_3)} = \frac{1/8}{(3/8)^3} = 64/27$$

Shogenji’s measure gives this problematic result because it fails to account for some of the relevant probabilistic information in the situation. It seems that what drives the intuition that  $\mathbf{E}$  is more coherent than  $\mathbf{E}'$  in the above case are the relations of positive relevance between the *individual* testimonies of the witnesses. In scenario one, there is agreement between each of the witnesses individually about two of the suspects, whereas in scenario two, there is only agreement between each of the witnesses individually about one of the suspects. Because Shogenji’s measure only accounts for *n*-wise dependencies, it misses out on the greater individual agreements in the first scenario and consequently mistakenly assigns the corresponding information sets equal degrees of coherence. This is an unappealing consequence of Shogenji’s generalized version of *Coh*.

<sup>3</sup>The quantitative details of this case were proposed to me by Jan Sprenger.

## 1.2 Problem of Irrelevant Addition

It seems quite clear that whenever one adds an irrelevant proposition to an otherwise coherent information set, the new set (resulting from the addition) must be less coherent than the first. Modifying an example from Bovens and Hartmann (2003), let  $S = \{R_1, R_2, R_3\}$ :

$R_1$ : The murderer had a French accent

$R_2$ : The murderer wore Coco Chanel shoes

$R_3$ : The murderer drove a Renault

Now say that we add  $R_4$  to  $S$ , so that we have  $S' = \{R_1, R_2, R_3, R_4\}$ :

$R_4$ : My aunt has the measles

In the absence of some story tying my Aunt's case of the measles in to the murder (in the light of which, the added proposition would no longer be irrelevant), it is clear that the degree of coherence of  $S = \{R_1, R_2, R_3\}$  is much greater than that of  $S' = \{R_1, R_2, R_3, R_4\}$

Shogenji's generalization gets this wrong however. In fact, for the generalized version of *Coh*, adding an irrelevant proposition to an information set necessarily does not change the degree of coherence of the information set at all. Let  $R_{n+1}$  be the irrelevant new addition to information set  $S = \{R_1, R_2, \dots, R_n\}$ ; then:

$$\begin{aligned} Coh(S \cup \{R_{n+1}\}) &= \frac{Pr(R_1 \wedge R_2 \wedge \dots \wedge R_{n+1})}{Pr(R_1) \times Pr(R_2) \times \dots \times Pr(R_{n+1})} \\ &= \frac{Pr(R_1 \wedge R_2 \wedge \dots \wedge R_n) \times Pr(R_{n+1})}{Pr(R_1) \times Pr(R_2) \times \dots \times Pr(R_n) \times Pr(R_{n+1})} \\ &= \frac{Pr(R_1 \wedge R_2 \wedge \dots \wedge R_n)}{Pr(R_1) \times Pr(R_2) \times \dots \times Pr(R_n)} = Coh(S) \end{aligned}$$

Thus, in general, adding an entirely irrelevant proposition to an information set must not affect that set's degree of coherence, according to *Coh*. This is an immensely unattractive consequence of Shogenji's generalized measure.

It is important to note that both the Depth Problem and the Problem of Irrelevant Addition necessarily involve information sets of cardinality  $n > 2$ . This is true in the former case given that there arises no Depth Problem for the case where  $n = 2$ ; in this case, the measure does dig as deeply into the information set as is possible - recall that this notion of coherence does not apply to singletons. And it is true in the latter case given that the Problem of Irrelevant Addition necessarily involves comparing the degrees of coherence of sets with different cardinalities; again, remembering that the relevant notion of coherence does not apply to singletons, at least one of these sets must have a cardinality  $n > 2$ . But given that these serious problems both involve sets of cardinality three or greater, one might naturally wonder if these problems have more to do with Shogenji's means of generalizing his measure than with the measure in full. In the following section, I

argue that this is the case and that there is an intuitive generalization of Shogenji's measure that also enables it to sidestep both of the serious problems presented above.

## 2 An Alternative Generalization of *Coh*

In both of the major problems above, the counterintuitive results given by Shogenji's generalized coherence measure may be seen as a consequence of the fact that it does not dig deeply enough into the probabilistic information about the relevant information sets. The Shogenji measure only considers  $n$ -wise (in)dependence, but this leaves the door open to counterexamples like those above where (1) an information set is intuitively more coherent than another information set due to probabilistic relations that are more intricate than those for which the Shogenji measure is able to account, or (2) a new (irrelevant) proposition is added to an  $n$ -membered information set which leaves the degree of relevance of the set on the  $n$ 'th level (but, as we shall see, *not* the degree of relevance on other levels) unchanged. A more intuitive generalization of Shogenji's measure might then take account of the degrees of positive relevance of a set *and* of all of its subsets (with cardinality  $n \geq 2$ ).

According to this suggestion, one should simply apply Shogenji's measure to all of the subsets (of cardinality  $n \geq 2$ ) of an information set and ultimately calculate the degree of coherence as some weighted average of these results. More formally, for any information set  $S = \{R_1, R_2, \dots, R_n\}$ , let  $s(S)$  be the logarithm of Shogenji's generalized coherence measure:<sup>4</sup>

$$s(S) =_{def} \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_2 \wedge \dots \wedge R_n)}{\text{Pr}(R_1) \times \text{Pr}(R_2) \times \dots \times \text{Pr}(R_n)} \right]$$

Then, for any natural number  $k$  that satisfies  $1 < k \leq n$ , we can define the  $k$ -wise coherence of an  $n$ -membered information set  $S$  as the mean of the values attained by applying the function  $s$  to all of the  $k$ -membered subsets of  $S$ . Let  $[S]^k$  denote the set of all subsets of  $S$  with cardinality  $k$  and recall that there are  $m = \binom{n}{k} = n!/k!(n-k)!$   $k$ -membered subsets of any  $n$ -membered set.

**Definition 1.**  $k$  – wise Coherence : Given a set  $S = \{R_1, \dots, R_n\}$  and an ordering  $\langle \tilde{S}_1, \dots, \tilde{S}_m \rangle$  of the members of  $[S]^k$ , the degree of  $k$ -wise coherence of  $S$  is given by the function

$$\mathcal{C}^k(S) =_{def} \frac{\sum_{i=1}^m s(\tilde{S}_i)}{m}$$

Once we have the degree of  $k$ -wise coherence calculated for all  $k$ , then the remaining question is how to weight these against each other in order to calculate an overall degree of coherence for  $S$ . In principle, any weighting scheme is open to testing. Leaving the specific weighting scheme unspecified, we have:

<sup>4</sup>I take the logarithm for normalization purposes, and because this move makes the task of combining various degrees of coherence (which is a necessary part of my generalized version of Shogenji's measure) more manageable.

**Definition 2.**  $\mathcal{C}$  : Given a set  $S = \{R_1, \dots, R_n\}$  and a weighting system that assigns the vector of positive weights  $\langle w_1, w_2, \dots, w_{n-1} \rangle$  to the  $k$ -wise degrees of coherence (where  $\sum_{i=1}^{n-1} w_i = 1$ ), the degree of coherence of  $S = \{R_1, \dots, R_n\}$  is given by the function

$$\mathcal{C}(S) =_{def} \sum_{i=1}^{n-1} w_i \times \mathcal{C}^{i+1}(S)$$

Notice that, were it not for the requirement that all weights be positive, Shogenji's generalized measure  $Coh$  could be seen as an instance of  $\mathcal{C}$  for which  $w_{n-1} = 1$  regardless of the cardinality of  $S$ . This point thus emphasizes the fact that, according to  $Coh$ , there is zero weight given to all degrees of  $k$ -wise coherence, when  $k \neq n$ . That is, Shogenji's generalization effectively ignores all sub-relations of coherence within the set.

In avoiding such a top-heavy, extreme weighting scheme, perhaps the most intuitive and obvious first attempt at an alternative lets each level of an information set count equally. That is, this weighting scheme takes the straight average of each  $k$ -wise degree of coherence:

**Definition 3.**  $\mathcal{C}$ (Straight Average) : Degree of coherence of  $S = \{R_1, \dots, R_n\}$  is given by the function

$$\mathcal{C}(S) =_{def} \frac{\sum_{k=2}^n \mathcal{C}^k(S)}{n-1}$$

Notice that  $\mathcal{C}$ (Straight Average) is just the special case of  $\mathcal{C}$  for which all weights are set equal to  $1/(n-1)$ . Of course, any number of weighting schemes other than Straight Average can be considered as well however. For the sake of this paper, I consider the following two additional alternatives:

**Definition 4.**  $\mathcal{C}$ (Deeper Decreasing) : Let the weighting scheme assign linearly decreasing weights to decreasing  $k$  in the following way:

$$w_i = \frac{i}{(n-1) + (n-2) + \dots + 1} = \frac{2i}{n(n-1)}$$

Then, degree of coherence of  $S = \{R_1, \dots, R_n\}$  is given by the function

$$\mathcal{C}(S) =_{def} \sum_{i=1}^{n-1} \frac{2i}{n(n-1)} \times \mathcal{C}^{i+1}(S) = \frac{\sum_{i=1}^{n-1} i \mathcal{C}^{i+1}(S)}{n(n-1)/2}$$

**Definition 5.**  $\mathcal{C}$ (Deeper Increasing) : Let the weighting scheme assign linearly increasing weights to decreasing  $k$  in the following way:

$$w_i = \frac{n-i}{(n-1) + (n-2) + \dots + 1} = \frac{2(n-i)}{n(n-1)}$$

Then, degree of coherence of  $S = \{R_1, \dots, R_n\}$  is given by the function

$$\mathcal{C}(S) =_{def} \sum_{i=1}^{n-1} \frac{2(n-i)}{n(n-1)} \times \mathcal{C}^{i+1}(S) = \frac{\sum_{i=1}^{n-1} (n-i) \mathcal{C}^{i+1}(S)}{n(n-1)/2}$$



Note that, for any of these specific versions of  $\mathcal{C}$ , if  $n = 2$ ,  $\mathcal{C}$  just reduces to (the Log-normalized version of) Shogenji's original measure. But for all information sets with cardinality  $n > 2$ ,  $\mathcal{C}$  offers a more intricate means for measuring the total degree of positive relevance of an information set. It is, so to speak, a *subset-sensitive* generalization of Shogenji's measure.

Lastly, given that this new generalization differs from Shogenji's original measure by making use of logarithms for normalization purposes, it is important to clarify the following: (any specific version of)  $\mathcal{C}$  is a function from a probability measure defined over the relevant propositions (the one's in the relevant information set  $S$ ) to a real number  $\mathcal{C}(S)$ , where (1)  $\mathcal{C}(S) > 0$  implies that  $S$  is coherent, (2)  $\mathcal{C}(S) < 0$  implies that  $S$  is incoherent, and (3)  $\mathcal{C}(S) = 0$  implies that  $S$  lacks any coherence (or incoherence) at all.

## 2.1 $\mathcal{C}$ and the Depth Problem

Any measure in the class  $\mathcal{C}$  easily avoids the Depth Problem. Applying  $\mathcal{C}$  to the previously problematic robbery case above, we get a much more satisfying result:

For  $\mathbf{E}$ , we get the following:

$$\begin{aligned}\mathcal{C}^2(\mathbf{E}) &= \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(W_1 \wedge W_2)}{\text{Pr}(W_1) \times \text{Pr}(W_2)} \right] + \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(W_1 \wedge W_3)}{\text{Pr}(W_1) \times \text{Pr}(W_3)} \right] \\ &\quad + \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(W_2 \wedge W_3)}{\text{Pr}(W_2) \times \text{Pr}(W_3)} \right] = \text{Log} \left[ \frac{1/4}{(3/8)^2} \right] = 0.250 \\ \mathcal{C}^3(\mathbf{E}) &= \text{Log} \left[ \frac{\text{Pr}(W_1 \wedge W_2 \wedge W_3)}{\text{Pr}(W_1) \times \text{Pr}(W_2) \times \text{Pr}(W_3)} \right] = \text{Log} \left[ \frac{1/8}{(3/8)^3} \right] = 0.375\end{aligned}$$

Thus:

$$\mathcal{C}(\mathbf{E}) = w_1 \times 0.250 + w_2 \times 0.375$$

Or (to take a specific weighting system as an example), in terms of  $\mathcal{C}$  (Straight Average), we get:

$$\mathcal{C}(\mathbf{E}) = \frac{0.250 + 0.375}{2} = 0.313$$

On the other hand, for  $\mathbf{E}'$ , we get the following:

$$\begin{aligned}\mathcal{C}^2(\mathbf{E}') &= \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(W'_1 \wedge W'_2)}{\text{Pr}(W'_1) \times \text{Pr}(W'_2)} \right] + \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(W'_1 \wedge W'_3)}{\text{Pr}(W'_1) \times \text{Pr}(W'_3)} \right] \\ &\quad + \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(W'_2 \wedge W'_3)}{\text{Pr}(W'_2) \times \text{Pr}(W'_3)} \right] = \text{Log} \left[ \frac{1/8}{(3/8)^2} \right] = -0.051\end{aligned}$$

$$\mathcal{C}^3(\mathbf{E}') = \text{Log} \left[ \frac{\text{Pr}(W'_1 \wedge W'_2 \wedge W'_3)}{\text{Pr}(W'_1) \times \text{Pr}(W'_2) \times \text{Pr}(W'_3)} \right] = \text{Log} \left[ \frac{1/8}{(3/8)^3} \right] = 0.375$$

Thus:

$$\mathcal{C}(\mathbf{E}') = w_1 \times -0.051 + w_2 \times 0.375$$

In terms of  $\mathcal{C}$ (Straight Average), we get:

$$\mathcal{C}(\mathbf{E}') = \frac{-0.051 + 0.375}{2} = 0.162$$

While  $\mathcal{C}$ (Straight Average) clearly matches our intuitions in this case, it is equally clear that any version of  $\mathcal{C}$  will do the job. Insofar as our weights must be positive,  $w_1 \times 0.250 + w_2 \times 0.375$  must always be greater than  $w_1 \times -0.051 + w_2 \times 0.375$  and thus  $\mathbf{E}$  must come out as more coherent than  $\mathbf{E}'$  - just as our intuitions mandate. Note that it is precisely the subset-sensitive nature of  $\mathcal{C}$  that allows it to give the intuitive result in this case in spite of the fact that Shogenji's original *Coh* does not. Looking at the above calculations, one notes that the 3-wise degree of coherence is equal between information sets; i.e.,  $\mathcal{C}^3(\mathbf{E}) = \mathcal{C}^3(\mathbf{E}')$ . This corresponds to the result given by Shogenji's generalized measure. However,  $\mathcal{C}$  goes beyond this result by also factoring in the differing 2-wise degrees of coherence between information sets:  $\mathcal{C}^2(\mathbf{E}) > \mathcal{C}^2(\mathbf{E}')$ .

That this result generalizes to information sets with cardinality  $n \geq 3$  is also easy to see. For any  $n$ -membered set  $S$ , Shogenji's generalized measure is equal to the  $n$ -wise coherence of  $S$ :  $\text{Coh}(S) = \mathcal{C}^n(S)$ . However, as defined above, for any such set,  $\mathcal{C}(S) = w_1 \times \mathcal{C}^2(S) + w_2 \times \mathcal{C}^3(S) + \dots + w_{n-1} \times \mathcal{C}^n(S)$ . Given that these weights are positive,  $k$ -wise degrees of coherence (where  $k \neq n$ ), which are ignored by  $\text{Coh}(S)$ , will affect the value of  $\mathcal{C}(S)$ .

## 2.2 $\mathcal{C}$ and the Problem of Irrelevant Addition

$\mathcal{C}$  also elegantly solves the specific version of the Problem of Irrelevant Addition in which an irrelevant proposition is added to an information set of cardinality two. As mentioned previously, when  $n = 2$ ,  $\mathcal{C}$  just reduces to Shogenji's original measure; thus:

$$\mathcal{C}(S) = \text{Log} \frac{\text{Pr}(R_1 \wedge R_2)}{\text{Pr}(R_1) \times \text{Pr}(R_2)}$$

With the addition of  $R_3$ ,  $\mathcal{C}$  gives the following results:

$$\begin{aligned} \mathcal{C}^2(S \cup \{R_3\}) &= \left\{ \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_2)}{\text{Pr}(R_1) \times \text{Pr}(R_2)} \right] + \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_3)}{\text{Pr}(R_1) \times \text{Pr}(R_3)} \right] \right. \\ &\quad \left. + \text{Log} \left[ \frac{\text{Pr}(R_2 \wedge R_3)}{\text{Pr}(R_2) \times \text{Pr}(R_3)} \right] \right\} / 3 = \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_2)}{\text{Pr}(R_1) \times \text{Pr}(R_2)} \right] \end{aligned}$$

$$\mathcal{C}^3(S \cup \{R_3\}) = \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_2 \wedge R_3)}{\text{Pr}(R_1) \times \text{Pr}(R_2) \times \text{Pr}(R_3)} \right] = \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_2)}{\text{Pr}(R_1) \times \text{Pr}(R_2)} \right]$$

Thus:

$$\mathcal{C}(S \cup \{R_3\}) = w_1 \times \frac{1}{3} \times \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_2)}{\text{Pr}(R_1) \times \text{Pr}(R_2)} \right] + w_2 \times \text{Log} \left[ \frac{\text{Pr}(R_1 \wedge R_2)}{\text{Pr}(R_1) \times \text{Pr}(R_2)} \right]$$

And these results imply (given that our weights must be positive and sum to 1) that  $|\mathcal{C}(S \cup \{R_3\})| < |\mathcal{C}(S)|$  for any irrelevant proposition  $R_3$ .

Notice that this result also implies that the addition of an irrelevant proposition to an otherwise *incoherent* set of information (one that is negatively coherent) must result in an *increase* in that set's degree of coherence. This makes intuitive sense too; the addition of an irrelevant piece of information to an information set should bring the degree of coherence of that set closer to the neutral point at which the set is judged to be lacking coherence altogether. For a set that is positively coherent, this means a decrease in degree of coherence; for a set that is negatively coherent, this means an increase in degree of coherence. In both cases, the result is a shift in degree of coherence toward the neutral value of zero.

The Problem of Irrelevant Addition proves to be more stubborn in cases where one's original information set  $S$  has a cardinality  $n > 2$ . In fact, only certain specific versions of  $\mathcal{C}$  are able to avoid this problem in such cases. As such, Irrelevant Addition provides a useful negative desideratum for our weighting scheme; i.e., any particular weighted version of  $\mathcal{C}$  that does not generally preclude this problem is not adequate for matching our intuitions about coherence. Thankfully, there exist such adequate weighting schemes that imply the desired result generally.

Pertaining to the three candidate weighting systems specified above (Straight Average, Deeper Decreasing, and Deeper Increasing), I show in **Appendix 1** that  $\mathcal{C}$ (Deeper Decreasing) does *not* generally avoid the Problem of Irrelevant Addition while  $\mathcal{C}$ (Straight Average) and  $\mathcal{C}$ (Deeper Increasing) *are* generally able to avoid this problem. This is perhaps not so surprising as  $\mathcal{C}$ (Deeper Decreasing) is, out of the three weighting systems considered here, the most similar to Shogenji's original generalization. Apparently then,  $\mathcal{C}$ (Deeper Decreasing) does not move far enough away from Shogenji's original generalized measure in order to avoid this problem.<sup>5</sup> On the other hand,  $\mathcal{C}$ (Straight Average) and  $\mathcal{C}$ (Deeper Increasing) stray further away from Shogenji's generalized version of his measure and thus are able to avoid the Problem of Irrelevant Addition. Thus, at least for  $\mathcal{C}$ (Straight Average) and  $\mathcal{C}$ (Deeper Increasing), adding an irrelevant proposition to an otherwise coherent set of information must result in a decrease in that set's degree of coherence. Consequently, these versions of  $\mathcal{C}$  are able to match our intuitions here.

<sup>5</sup>This is, of course, not to say that there could not be other weighting systems that both decrease the relative weight given to deeper levels of an information set and avoid the Problem of Irrelevant Addition. Thanks to Luc Bovens for suggesting this intuitive account of why Deeper Decreasing still falls prey to the Problem of Irrelevant Addition.

### 3 The Way Ahead

In this paper, I have shown that the two most worrisome objections to Shogenji's formal account of coherence necessarily involve information sets of cardinality  $n \geq 3$ . Given this, I surmised that the problem with Shogenji's measure has more to do with his means of generalizing the measure than with the measure itself. I defended this claim by offering  $\mathcal{C}$  as an alternative generalization of Shogenji's measure.  $\mathcal{C}$  retains the intuitive merits of Shogenji's measure while - under certain weighting systems - avoiding both of the relevant problems that befall it. In the light of all of this then, I suggest that there is new hope for Shogenji's measure; Shogenji's early and influential attempt at formalizing coherence, when generalized in a subset-sensitive way, is able to clear the objections that plagued *Coh*.

So where do we go from here? Although I have defended  $\mathcal{C}$  against some important objections in this paper, there is still much work to be done if  $\mathcal{C}$  is to be accepted as an accurate formal analysis of our concept of coherence. Perhaps most importantly, the respective validity of various weighting systems needs to be explored. As mentioned above, Straight Average, Deeper Decreasing, and Deeper Increasing are of course not the only options. There are any number of alternative weighting systems, many of which are non-arithmetic. However, instead of viewing this as a weakness of  $\mathcal{C}$ , I take it that this consideration points to an expansive arena for further research.

We have already seen that the philosophical methods of intuition-pumping through counterexamples and thought experiments will continue to be useful in this arena - recall my suggestion above that the Problem of Irrelevant Addition act as a negative desideratum for judging weighting systems; recall also that we successfully ruled out one particular weighting system, Deeper Decreasing, via this class of counterexamples. However, where such intuitions begin to cloud, more empirical, experimental methods should also be brought to bare upon this issue. For example, one promising psychological study could explore the comparative levels of psychological validity for a host of differently weighted versions of  $\mathcal{C}$ . Presumably, such a study would proceed by comparing test participant's actual intuitions about cases involving coherence with the resulting degrees of coherence that fall out of each particular measure. In this task,  $\mathcal{C}$ 's current lack of a specific weighting system, along with the great variety of possible weighting systems to apply, would be far from a weakness. Instead, this feature gives  $\mathcal{C}$  a great freedom in its ability to match up with real human intuitions about coherence. In other words, the higher the number of plausible, candidate coherence measures that result from filling in the weighting details of  $\mathcal{C}$ , the higher the chances that one of these will be psychologically valid.

Another crucial topic for further research has to do with truth-conduciveness. There is a simple and significant sense in which Shogenji's original notion of coherence promotes higher probabilities of a set's being true (and so  $\mathcal{C}$  retains this appealing feature when applied to sets of cardinality two); however, the additional details involved in making  $\mathcal{C}$  subset-sensitive also obscure its relation to truth (in cases where an information set has cardinality  $n > 2$ ). Thus, more formal work needs to be done in order to spell out whether there are any interesting connec-

tions between  $\mathcal{C}$  and the probability of an information set being true.

Clearly then, there is much important work yet to be done. Nonetheless, my hope is that, in this paper, I have provided sufficient motivation for this further research by showing that Shogenji's measure, in the alternatively generalized form  $\mathcal{C}$ , is still very much alive as a candidate for a formal account of coherence. In the light of this paper then, I suggest that Bayesians interested in coherence return to their humble beginnings and reconsider Shogenji's notion of coherence as positive relevance.

## Appendix 1.

Here, I investigate whether  $\mathcal{C}$ (Straight Average),  $\mathcal{C}$ (Deeper Decreasing), and  $\mathcal{C}$ (Deeper Increasing) fall prey to the general version of the Problem of Irrelevant Addition (for any  $n$ ).

First note that, for any version of  $\mathcal{C}$ , when an irrelevant proposition  $I$  is added to information set  $S$  giving  $S' = S \cup \{I\}$ ,  $\mathcal{C}(S')$  is a function of the same terms as  $\mathcal{C}(S)$  (given that the new proposition - being irrelevant - may be systematically factored out of all terms in the new calculation). In addition, in accounting for the weights and averaging the degree of coherence at each level of the set, each of these terms is multiplied by a factor related to the previous set of weights as follows - where  $n$  refers to the cardinality of the original information set  $S$ , and where  $w'_i$  is the new weight given to the  $k(= i + 1)$ -wise coherence of  $S'$ ,  $\mathcal{C}^k(S')$ :

$$\frac{w_i}{\binom{n}{i+1}} \rightarrow \frac{w'_{i+1}}{\binom{n+1}{i+2}} + \frac{w'_i}{\binom{n+1}{i+1}}$$

For example, let  $S = \{A, B, C\}$ . Then:

$$\begin{aligned} \mathcal{C}(S) = & w_2 \times \text{Log} \frac{Pr(A \wedge B \wedge C)}{Pr(A) \times Pr(B) \times Pr(C)} + \frac{w_1}{3} \times \\ & \left[ \text{Log} \frac{Pr(R_1 \wedge R_2)}{Pr(R_1) \times Pr(R_2)} + \text{Log} \frac{Pr(R_1 \wedge R_3)}{Pr(R_1) \times Pr(R_3)} + \text{Log} \frac{Pr(R_2 \wedge R_3)}{Pr(R_2) \times Pr(R_3)} \right] \end{aligned}$$

Adding irrelevant  $I$  to  $S$  results in  $S' = \{A, B, C, I\}$ . In this case, remembering that  $I$  is by definition irrelevant to all subsets of  $S$ , we get:

$$\begin{aligned} \mathcal{C}(S') = & \left[ w'_3 + \frac{w'_2}{4} \right] \times \text{Log} \frac{Pr(A \wedge B \wedge C)}{Pr(A) \times Pr(B) \times Pr(C)} + \left[ \frac{w'_2}{4} + \frac{w'_1}{6} \right] \times \\ & \left[ \text{Log} \frac{Pr(R_1 \wedge R_2)}{Pr(R_1) \times Pr(R_2)} + \text{Log} \frac{Pr(R_1 \wedge R_3)}{Pr(R_1) \times Pr(R_3)} + \text{Log} \frac{Pr(R_2 \wedge R_3)}{Pr(R_2) \times Pr(R_3)} \right] \end{aligned}$$

$\mathcal{C}(S')$  is a function of the same terms as  $\mathcal{C}(S)$ ; however, these terms are multiplied by different factors according to the previously stated pattern:

$$\begin{aligned} \frac{w_2}{\binom{3}{3}} = w_2 & \rightarrow \frac{w'_3}{\binom{4}{4}} + \frac{w'_2}{\binom{4}{3}} = w'_3 + \frac{w'_2}{4} \\ \frac{w_1}{\binom{3}{2}} = \frac{w_1}{3} & \rightarrow \frac{w'_2}{\binom{4}{3}} + \frac{w'_1}{\binom{4}{2}} = \frac{w'_2}{4} + \frac{w'_1}{6} \end{aligned}$$

Given this relation between multipliers, we can prove that  $\mathcal{C}(S) > \mathcal{C}(S')$  for a particular weighting scheme - and thus that the version of  $\mathcal{C}$  resulting from applying that weighting scheme avoids the Problem of Irrelevant Addition - by showing, for that scheme, that it must be the case that:

$$\frac{w_i}{\binom{n}{i+1}} > \frac{w'_{i+1}}{\binom{n+1}{i+2}} + \frac{w'_i}{\binom{n+1}{i+1}}$$

Reducing this inequality to a simpler form, this amounts to showing that the following must hold, given a particular weighting scheme:

$$w_i > \frac{w'_{i+1} \times (i+2) + w'_i \times (n-i)}{n+1}$$

Similarly, we can show that a weighting scheme falls prey to the Problem of Irrelevant Addition by showing that, given that weighting scheme and particular values for  $n$  and  $i$ , the reverse inequality holds. So long as this may occur for some values of  $n$  and  $i$ , then counterexamples can be thought of wherein an information set is sufficiently  $k-1 = i$ -wise coherent to make the overall coherence of  $S$  less than that of  $S'$ . These are the two approaches that I take in proving the following three theorems.

1.  $\mathcal{C}$ (Straight Average) avoids the Problem of Irrelevant Addition.

In general, given  $\mathcal{C}$ (Straight Average), the following hold:

$$w_i = \frac{1}{n-1}, w'_i = \frac{1}{n}, w'_{i+1} = \frac{1}{n}$$

Thus:

$$\frac{w'_{i+1} \times (i+2) + w'_i \times (n-i)}{n+1} = \frac{n+2}{n(n+1)}$$

And it must be the case that

$$w_i = \frac{1}{n-1} > \frac{n+2}{n(n+1)}$$

as made obvious by the following analysis of this inequality:

$$1 > \frac{(n+2)(n-1)}{n(n+1)} \Rightarrow (n+2)(n-1) < n(n+1) \Rightarrow n^2 + n - 2 < n^2 + n$$

□

2.  $\mathcal{C}$ (Deeper Increasing) avoids the Problem of Irrelevant Addition.

In general, given  $\mathcal{C}$ (Deeper Increasing), the following hold:

$$w_i = \frac{2(n-i)}{n(n-1)}, w'_i = \frac{2(n-i+1)}{n(n+1)}, w'_{i+1} = \frac{2(n-i)}{n(n+1)}$$

Thus:

$$\begin{aligned} & \frac{w'_{i+1} \times (i+2) + w'_i \times (n-i)}{n+1} \\ &= \frac{2(n-i) \times (i+2) + 2(n-i+1) \times (n-i)}{n(n+1)^2} \\ &= \frac{2(n-i)(n+3)}{n(n+1)^2} \end{aligned}$$

And it must be the case that

$$w_i = \frac{2(n-i)}{n(n-1)} > \frac{2(n-i)(n+3)}{n(n+1)^2}$$

as made obvious by the following analysis of this inequality:

$$\frac{1}{n-1} > \frac{n+3}{(n+1)^2} \Rightarrow (n+1)^2 > (n+3)(n-1) \Rightarrow n^2 + 2n + 1 > n^2 + 2n - 3$$

□

3.  $\mathcal{C}$ (Deeper Decreasing) *does not avoid the Problem of Irrelevant Addition.*

In general, given  $\mathcal{C}$ (Deeper Decreasing), the following hold:

$$w_i = \frac{2i}{n(n-1)}, w'_i = \frac{2i}{n(n+1)}, w'_{i+1} = \frac{2(i+1)}{n(n+1)}$$

Thus:

$$\begin{aligned} & \frac{w'_{i+1} \times (i+2) + w'_i \times (n-i)}{n+1} = \frac{2(i+1)(i+2) + 2i(n-i)}{n(n+1)^2} \\ &= \frac{2(2+i(n+3))}{n(n+1)^2} \end{aligned}$$

It is *not* always the case, for any  $i$  and any  $n$ , that this term must be less than  $w_i$ . A graph of  $w_1$  and  $2(2+(n+3))/n(n+1)^2$  (i.e., setting  $i=1$ ) as functions of  $n$  shows this to be true (**Figure 1**). When  $n < 2$ ,  $w_1 > 2(2+(n+3))/n(n+1)^2$ ; however, for all  $n > 2$ , the opposite inequality holds:  $w_1 < 2(2+(n+3))/n(n+1)^2$ .

□



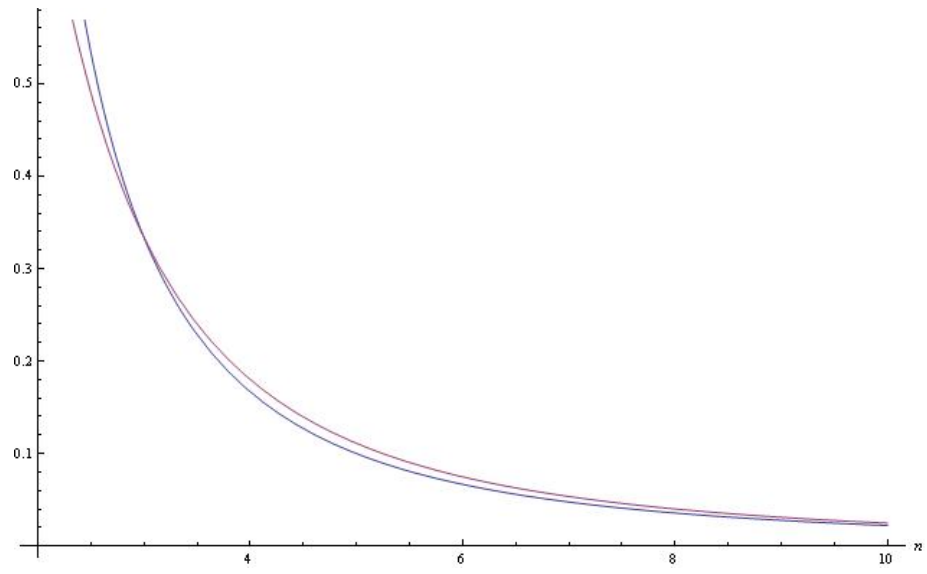


Figure 1.

## References

- Akiba, K. (2000). Shogenji's probabilistic measure of coherence is incoherent. *Analysis*, 60(4):356–359.
- Bovens, L. and Hartmann, S. (2003). *Bayesian Epistemology*. Oxford University Press, New York, NY.
- Douven, I. and Meijs, W. (2007a). Measuring coherence. *Synthese*, 156:405–425.
- Douven, I. and Meijs, W. (2007b). On the alleged impossibility of coherence. *Synthese*, 157:347–360.
- Fitelson, B. (2003). A probabilistic theory of coherence. *Analysis*, 63(3):194–199.
- Fitelson, B. (2004). Two technical corrections to my coherence measure. Unpublished. Available online at <http://fitelson.org/coherence2.pdf>.
- Glass, D. H. (2005). Problems with priors in probabilistic measures of coherence. *Erkenntnis*, 63:375–385.
- Meijs, W. (2005). *Probabilistic Measures of Coherence*. PhD thesis, Erasmus Universiteit Rotterdam.
- Meijs, W. (2006). Coherence as generalized logical equivalence. *Erkenntnis*, 64:231–252.
- Olsson, E. J. (2002). What is the problem of coherence and truth? *Journal of Philosophy*, 94:246–272.
- Olsson, E. J. (2005). *Against Coherence: Truth, Probability, and Justification*. Oxford University Press, New York, NY.
- Schupbach, J. N. (2008). On the alleged impossibility of bayesian coherentism. *Philosophical Studies*, 141(3).
- Shogenji, T. (1999). Is coherence truth-conducive? *Analysis*, 59(4):338–345.
- Shogenji, T. (2001). Reply to akiba on the probabilistic measure of coherence. *Analysis*, 61(2):147–150.
- Wheeler, G. (2009). Focused correlation and confirmation. *British Journal for the Philosophy of Science*, 60(1):79–100.