Structured probabilistic coherence and the usual counterexamples to probabilistic measures of coherence

Contents

1 Introduction & motivations

The notion of coherence is often used in many philosophical, especially epistemological, discussions (for instance, in discussions about the truth-conduciveness of coherence). When we talk about the coherence of a set of propositions or about the coherence of a story, we seem to refer to how well their individual pieces fit together. How are we to understand and apply this notion systematically, though? In particular, we will be interested in probabilistic explications of this notion, as Bayesian epistemology strives to be a general epistemological project and as such it should be able to accommodate coherence-oriented considerations.

There is also a more practical reason to develop a better understanding of the notion: a plausible measure of coherence could be used to better evaluate the quality of some stories or narrations. For example in the legal context we would like to be able to assess the quality of a testimony in the court of law.

Multiple probabilistic explications of coherence have been proposed (Douven & Meijs, 2007; Fitelson, 2003; Glass, 2002; Meijs & Douven, 2007; Olsson, 2001; Roche, 2013; Shogenji, 1999). However, clear general principles to choose between them are hard to come by. One paper where some such principles have been formulated is (Schippers, 2014), where a list of seemingly plausible adequacy conditions for a coherence measure is proposed and shown to be inconsistent to argue for pluralism about the notion of coherence. However, some of those requirements are quite non-trivial.[1]

The general point here is not that the approach taken in (Schippers, 2014) is flawed, but rather that the task of formulating general principles for coherence is a challenge, and that no clear list of such uncontroversial desiderata is on the horizon.

One approach to obtaining some clarity on which abstract conditions are plausible is looking at various thought experiments in which our intuitions about what the coherence scores should be (at least comparatively) are more robust than direct assessment of general requirements. In fact, looking at examples is what the main stream of literature on probabilistic coherence focused on, and each probabilistic measure of coherence faces a selection of seemingly intuitive counterexamples.

We decided to work with this methodology. We first gathered key examples that occur in the literature, represented them in terms of Bayesian networks, and developed **R** scripts calculating all coherence scores for the Bayesian networks at play, pushing further the results obtained by Koscholke (2016).[2] Then we reflected on the results, noticing that one weakness of the measures is that they pay little attention to the underlying structure of a given narration in the calculation of its coherence.

Inspired by this observation, we formulate our own proposal, which diverges from the known purely probabilistic measures of coherence in three important respects: (i) It is not a function of a probability measure and a set of propositions alone, because it is also sensitive to the selection and direction of edges in a Bayesian network representing an agent's credal state. (ii) Unlike in the case of quite a few coherence measures, it is sensitive to the weakest links in the narration. (iii) It is not obtained by simply averaging confirmation levels between all possible combinations of elements.

We described this approach in a more detailed introduction to this measure [ANONYMIZED], which explains the method and some of the theoretical decisions that we have made, and show how it works using a Bayesian network developed for the well-known Sally Clark case (Fenton & Neil, 2018). The goal of the current paper is to discuss a range of philosophical counterexamples to the existing probabilistic measures of coherence and evaluate the performance of our approach using those as a benchmark, arguing that it performs much better than the existing ones.

Accordingly, in Section 2 we introduce all the coherence measures, including the key motivations for and a pseudo-code description of our measure. In Section 3 we describe the thought experiments meant

[1] Let us illustrate this. The (Dependence) condition formulated there requires that the coherence score of a set of propositions is above (below) the neutral score if for all pairs of non-empty subsets the posterior of an element of a pair conditional on the other element is higher than the prior of the former. This makes some of the features of the coherence measure dependent on the priors, and whether it should be so is not obvious. On the other hand, (Agreement) is formulated in terms of conditional probabilities between such pairs. If on a given measure P all conditional probabilities (between pairs already mentioned) are higher than on P', the coherence of a set given P should be higher than given P'. The (Equivalence) requirement is that any finite set of logically equivalent propositions should be maximally coherent. This is suspicious, as the set $\{0 = 1, 2, = 5\}$ is a set of equivalent propositions (with sufficiently strong notion of logical equivalence in the background), but we would intuitively hesitate to say it's maximally coherent.

[2] The whole work has been made possible by all those who contributed to the development of **R** language, and Marco Scutari, the author of **bnlearn** package, who was kind enough to extend his package with additional features upon our requests (Scutari & Denis, 2015).

as counterexamples to coherence measures, their corresponding desiderata and their status on various coherence measures, including ours. The order of the discussion of any given example is straightforward: we first explain what the situation we are to consider is, what the intuitive desiderata related to it are supposed to be, how the situation is represented by means of a Bayesian network(s), and what happens when we apply all coherence measures. We end with Section **??** in which we compare all of the results and draw some general conclusions.

2 Probabilistic coherence measures and structured coherence

Quite a few different measures of coherence have been proposed in the literature. Two early proposals are:

- Shogenji's **deviation from independence** (Shogenji, 1999), is defined as the ratio between the probability of the conjunction of all claims, and the probability that the conjunction would get if all its conjuncts were probabilistically independent (scaling from 0 to ∞ with neutral point 1):

$$\mathscr{C}_S(S) = \frac{P(\bigwedge S)}{\prod_{i=1}^{|S|}\{P(S_i)|i \in S\}} \qquad \text{(Shogenji)}$$

This measure was later generalized by Meijs & Douven (2007). According to this approach, (Shogenji) is applied not only to the whole set of propositions, but to each non-empty non-singleton subset of the set, and the final value is defined as the average of all sub-values thus obtained.

- **Relative overlap** (Glass, 2002; Olsson, 2001), is defined as the ratio between the intersection of all propositions and their union (scaling from -1 to 1 with no clear neutral point):

$$\mathscr{C}_O(S) = \frac{P(\bigwedge S)}{P(\bigvee S)} \qquad \text{(Olsson)}$$

It has also been generalized in a way analogous to the one used in the generalization of the Shogenji's measure (Meijs & Douven, 2007).

Both of these approaches are susceptible to various objections and counterexamples (Akiba, 2000; Bovens & Hartmann, 2004; Crupi, Tentori, & Gonzalez, 2007; Koscholke, 2016; Merricks, 1995; Schippers & Koscholke, 2019; Shogenji, 1999, 2001, 2006; Siebel, 2004, 2006). To overcome them, more recent works proposed **average mutual support** measures, starting with (Fitelson, 2003). The general recipe for such measures is as follows.

- Given that S is a set whose coherence is to be measured, let P indicate the set of all ordered pairs of non-empty, disjoint subsets of S.
- First, define a confirmation function (of a hypothesis H by evidence E): $\mathsf{conf}(H,E)$.
- For each pair $\langle X,Y\rangle \in P$, calculate $\mathsf{conf}(\bigwedge X, \bigwedge Y)$, where $\bigwedge X$ is the conjunction of all the elements of X (and $\bigwedge Y$ is to be understood analogously).
- Take the mean of all the results:

$$\mathscr{C}(S) = \mathsf{mean}\left(\left\{\mathsf{conf}(\bigwedge X_i, \bigwedge Y_i)|\langle X_i, Y_i\rangle \in P\right\}\right).$$

Different measures of coherence result from different choices of a confirmation measure. Here are the key candidates present in the literature:

- Fitelson (2003) uses the following confirmation function (the resulting coherence measure ranges from -1 to 1 with neutral point at 0):

$$F(H,E) = \begin{cases} 1 & E \models H, E \not\models \bot \\ -1 & E \models \neg H \\ \frac{P(E|H)-P(E|\neg H)}{P(E|H)+P(E|\neg H)} & \text{o/w} \end{cases}$$

$$\mathscr{C}_F(S) = \mathsf{mean}\left(\left\{F(\bigwedge X_i, \bigwedge Y_i)|\langle X_i, Y_i\rangle \in P\right\}\right) \qquad \text{(Fitelson)}$$

- Douven & Meijs (2007) use the difference confirmation measure (with coherence ranging from -1 to 1 with neutral point at 0):

$$D(H,E) = P(H|E) - P(H)$$

$$\mathscr{C}_{DM}(S) = \mathsf{mean}\left(\left\{D(\bigwedge X_i, \bigwedge Y_i)|\langle X_i, Y_i\rangle \in P\right\}\right) \qquad \text{(DM)}$$

- Roche (2013) uses the absolute confirmation measure (the resulting coherence measure ranges from 0 to 1 with neutral point at 0.5):

$$A(H,E) = \begin{cases} 1 & E \models H, E \not\models \bot \\ 0 & E \models \neg H \\ P(H|E) & \text{o/w} \end{cases}$$

$$\mathscr{C}_R(S) = \mathsf{mean}\left(\left\{A(\bigwedge X_i, \bigwedge Y_i) | \langle X_i, Y_i\rangle \in P\right\}\right) \qquad \text{(Roche)}$$

Mind your head: different measures use different scales and have different neutral points (values taken for any set of probabilistically independent propositions; not all measures have neutral points). This is worth keeping in mind when it comes to various desiderata that we will discuss.

As we already mentioned in the introduction, formulating abstract formal requirements for a coherence measure and investigating whether a given coherence measure satisfies them has not resulted in an agreement. For this reason, we follow another path, which has dominated the literature on the topic. We look at how the measures behave in test scenarios. Many putative scenarios were put forward as counterexamples. They usually have the form of a few propositions formulated in natural language, such that intuitive judgments of coherence involved and the formal coherence calculations seem to diverge (Akiba, 2000; Bovens & Hartmann, 2004; Koscholke, 2016; Meijs & Douven, 2007, 2007; Merricks, 1995; Schippers & Koscholke, 2019, 2019; Shogenji, 1999, 2001, 2006; Siebel, 2004, 2006). We will focus on these examples in what follows. To spoil the experience, let us already point out that the probabilistic measures we introduced above do not seem to handle these examples very well (read on for details).

Inspired by these failures, in [REFERENCE SUPRESSED FOR ANONYMITY] we proposed to take a different perspective. Putting the earliest measures aside (they were problematic for various reasons), we noticed that the problems with the average mutual support measures stem from the fact that the coherence score is an average confirmation score for all possible combinations of the parts of a narration. Therefore we proposed to take a more fine-grained account. First, we represented an agent's belief state by means of a Bayesian network, which comprises not only a probabilistic measure but additional structural information, Then we used this structural information in our definition of coherence, so that only those directions of support are considered which in fact are indicated by the structure of the agent's belief state.

While we refer the reader to a more extensive treatment in [REFERENCE SUPRESSED FOR ANONYMITY], we now briefly discuss the main idea behind it. A Bayesian network represents agent's probabilistic belief state with respect to the relevant nodes. Some of them are distinguished as fixed narration nodes—the agent holds definite beliefs about which states of these nodes occur.

Each parented node in the BN receives its expected confirmation score (ECS). It is calculated by looking at all combinations of its states and states of its parents not excluded by agents' fixed beliefs. For each of these combinations, the confirmation score between the parents' states and the child state is calculated (in the pseudo-code, we use confirmation measure Z, in further calculations we also use measures LR and L for comparison)[3]. Then, a weighted average of these scores is obtained, where weights are the probabilities of the combinations of parents' states obtained by updating the BN with the fixed states of the fixed narration nodes. The final coherence score is either the mean of the ecs scores, if all of them are positive, or it is a weighted average of their mean and their minimum, $(1-|\mathsf{min}(\mathsf{ecs})|)\times\mathsf{mean}(\mathsf{ecs})+|\mathsf{min}(\mathsf{ecs})|\times\mathsf{min}(\mathsf{ecs})$, otherwise.[4]

[3] DEFINITIONS

[4] We have developed **R** code calculating this and other measures to handle calculations that will be discussed further on, the code with documentation is available at ANONYMIZED.

```
FUNCTION parents_child_possible_states(parents,child)
    IF child included in narration THEN
        consequentStates <- the unique state of child as reported in the narration
    ELSE
        consequentStates <- all possible states of child
    FOR EACH parent in parents
        IF parent included in narration THEN
            parentStates[parent] <- the unique state of parent as reported in the narration
        ELSE
            parentStates[parent] <- all possible states of parent
    parentsStates <- all combinations of parentStates
    variants <- list of all possible combinations of consequentStates and parentsStates
    RETURN variants


FUNCTION coherence_from_ecs(ecs)
    IF min(ecs) <= 0 THEN
        RETURN mean(ecs) * (min(ecs)+1) - min(ecs)min(ecs)
        #this is equivalent to (1- |min(ecs)|) * mean(ecs) + |min(ecs)| * min(ecs)
    ELSE
        RETURN mean(ecs)

FUNCTION structured_coherence(BN,fixedNodes,fixedStates)
    parentedNodes <- vector of non-root nodes in BN
    ecsList <- []
    FOR EACH parentedNode IN parentedNodes
        variants <- parents_child_possible_states(parents,parentedNode)
        variants_count <- length(variants)
        sumParentsNarr <- 0
        FOR EACH variant IN variants
            childPrior <- prior probability of the child state in variant
            childPosterior <- posterior probability of the child state in variant,
                        obtained by updating on the parents states
                        from this variant
            parentsNarr[variant] <- joint probability of
                        the parents states in variant in BN updated with fixedStates of fixedNodes
            sumParentsNarr <- sumParentsNarr +  parentsNarr[variant]
            z[variant] <- z_confirmation_measure(childPrior, childPosterior)
        ecs <- 0  #expected confirmation score
        FOR EACH variant IN variants
            IF parentsNarr[variant] > 0 THEN
                weight <- parentsNarr[variant]/sumParentsNarr
            ELSE
                weight <- 1/variants_count
            zScaled <- z[variant] * weight
            ecs <- ecs + zScaled
        ecsList.add(ecs)
    RETURN coherence_from_ecs(ecsList)
```

Having introduced the coherence measures at play, let us now move to the key counterexamples discussed in the literature.

3 Challenges and their treatment

For each of the counterexamples, we first explain what it is and what the connected desiderata are. Then we represent it as a Bayesian network, and finally we use our **R** scripts to calculate coherence scores that the coherence measures included in the previous section (including ours) yield for a given example and whether the desiderata are satisfied. Here are the counterexamples put forward against various coherence measures in the literature. We ignored only a few where both we didn't share the authors' intuitions and the examples were not picked up in further discussion in the literature.

3.1 Penguins

The scenario. A challenge discussed in (Bovens & Hartmann, 2004, p. 50) and (Meijs & Douven, 2007) consists of the propositions (instead of *letters* or *abbreviations*, we'll talk about *nodes*, as these will be used later on in Bayesian networks) displayed in Table 1.

node	content
B	Tweety is a bird.
G	Tweety is a grounded animal.
P	Tweety is a penguin.

Table 1: Propositions in the Penguins scenario

Desiderata. It seems that the set {B,G}, which doesn't contain the information about Tweety being a penguin, should be less coherent than the one that does contain this information: {B,G,P}.

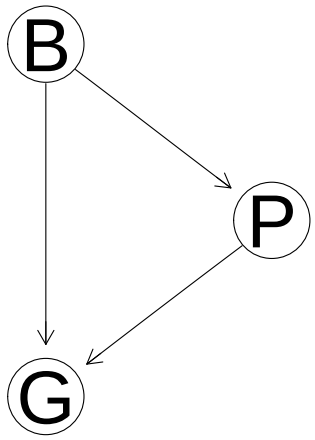
(BG<BGP) {B,G} should be less coherent than {B,G,P}.

Another intuition about this scenario (Schippers & Koscholke, 2019) is that when you consider a set which says that Tweety is both a bird and a penguin: {B,P}, adding proposition about not flying (G) shouldn't increase the coherence of the set as much as moving from {B,G} to {B,G, P}. It's a well-known fact that penguins don't fly and by adding G explicitly to the set, one wouldn't gain as much information. However, as G is not a logical consequence of P, it can be argued that {B,P} and {B,P,G} represent different information sets, and so some difference in their coherence is to be expected.

(BG ≪ BP≤BGP) {B,P} should be notably above {B,G}, and less than {B,P,G}.

Formally, we'll require that the absolute difference between BG and BP be greater than .1 (the exact placement of the threshold doesn't make a huge difference, unless it's at an unintuitive value below .01) and that {B,G} ≤ {B,P,G}.

Bayesian network. We used the distribution used in the original formulation to build a BN corresponding to the narrations at play (Fig. 1).[5]



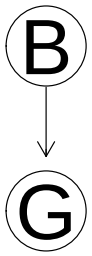
B	Pr
1	0.5
0	0.5

P	B	
	1	0
1	0.02	0
0	0.98	1

	B	P	
G			Pr
1	1	1	1.00
0	1	1	0.00
1	0	1	0.00
0	0	1	1.00
1	1	0	0.00
0	1	0	1.00
1	0	0	0.98
0	0	0	0.02

Figure 1: Bayesian network for the Penuins problem.

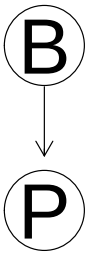
[5]Not without concerns. There are around 18 000 species of birds, and around 60 of them are flightless. We couldn't find information about counts, but it seems the probability of being a penguin if one is grounded is overestimated by philosophers. Also, there are many things that are not grounded but are not birds, mostly insects, and there's plenty of them. We did spend some time coming up with plausible ranges of probabilities to correct for such factors, and none of them actually makes a difference to the main point. So, for the sake of simplicity, we leave the original unrealistic distribution in our discussion.



B	Pr
1	0.5
0	0.5

G	B	
	1	0
1	0.02	0.98
0	0.98	0.02

Figure 2: Bayesian network for the BG scenario.



B	Pr
1	0.5
0	0.5

P	B	
	1	0
1	0.02	0
0	0.98	1

Figure 3: Bayesian network for the BP scenario.

Results. Now, let's calculate the coherence scores (Table 2) and see if the desiderata are satisfied (Table 3). The measures are: Olsson-Glass, generalized Olsson-Glass, Shogenji, generalized Shogenji, Douven-Meijs, Roche, Fitelson, Structured with Z, LR, and L used as a confirmation measure.

	OG	OGGen	Sh	ShGen	DM	R	Fi	SZ	SLR	SL
Penguins: BGP 111	0.01	0.015	4.00	2.01	0.255	0.51	0.453	0.505	Inf	0.669
Penguins: BG 11	0.01	0.010	0.04	0.04	-0.480	0.02	-0.960	-0.960	0.02	-0.960
Penguins: BP 11	0.02	0.020	2.00	2.00	0.255	0.51	0.669	0.010	2.02	0.338

Table 2: Coherence scores for the Penguins scenario (rounded). Note how LR might result in Inf if a conditional probability of 1 at an arrow used in the calculations is involved.

	OG	OGGen	Sh	ShGen	DM	R	Fi	SZ	SLR	SL
Penguins: BG<BGP	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
Penguins: BG<< BP< BGP	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE

Table 3: Desiderata satisfaction for the Penguins scenario.

3.2 Dunnit

The scenario. Another challenge, introduced by Merricks (1995) goes as follows: Mr. Dunnit is a suspect in the murder case. Detectives first obtained the body of evidence specified in Table 4.

node	content
I	Witnesses claim to have seen Dunnit do it (incriminating testimony).
M	Dunnit had a motive for the murder.
W	A credible witness claims to have seen Dunnit two hundred miles from the scene of the crime at the time of the murder.

Table 4: Initial evidence in the Dunnit scenario.

In light of this information they try to assess whether Dunnit is responsible for the crime (Table 5.

node	content
G	Dunnit is guilty.

Table 5: The guilt statement in the Dunnit scenario.

Now, suppose the detectives learn Dunnit has a twin brother (Table 6).

node	content
Tw	Dunnit has an identical twin which was seen by the credible witness two hundred miles from the scene of the crime during the murder.

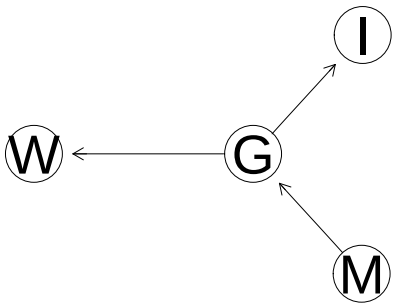
Table 6: New evidence in the Dunnit scenario.

What are our intuitions when we compare the coherence of $\{\mathsf{I},\mathsf{M},\mathsf{W},\mathsf{G}\}$ with the coherence of $\{\mathsf{I},\mathsf{M},\mathsf{W},\mathsf{G},\mathsf{Tw}\}$?

Desideratum. It seems that adding proposition about a twin should increase the coherence of the set.

(Dunnit<Twin) $\{\mathsf{I},\mathsf{M},\mathsf{W},\mathsf{G}\}$ should be less coherent than $\{\mathsf{I},\mathsf{M},\mathsf{W},\mathsf{G},\mathsf{Tw}\}$.

Bayesian networks. Here, we deal with two separate BNs. One, before the Twin node is even considered (Figure 4), and one with the Twin node (Figure 5). The CPTs for the no-twin version are in agreement with those in the ones in the Twin case. Since the original example didn't specify exact probabilities, we came up with some plausible values.



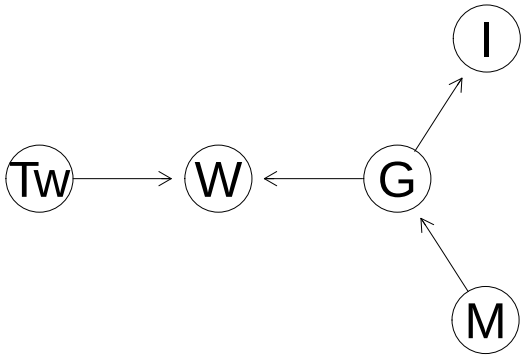
M	Pr
1	0.4
0	0.6

G	M	
	1	0
1	0.05	0.005
0	0.95	0.995

I	G	
	1	0
1	0.8	0.005
0	0.2	0.995

W	G	
	1	0
1	0.012	0.207
0	0.988	0.793

Figure 4: Twin-less BN for the Dunnit problem.



G	Tw		
W			Pr
1	1	1	0.200
0	1	1	0.800
1	0	1	0.400
0	0	1	0.600
1	1	0	0.005
0	1	0	0.995
1	0	0	0.200
0	0	0	0.800

Figure 5: BN for the Dunnit problem. The key difference for the twin version lies in the construction of the CPT for W. The table gives conditional probabilities for W given various joint states of Tw and G.

Results. Coherence calculations result in Table 7 and how they fare with respect to the desideratum is displayed in Table 8.

	OG	OGGen	Sh	ShGen	DM	R	Fi	SZ	SLR	SL
Dunnit: MGWI 1111	0	0.087	4.294	11.012	0.167	0.266	0.169	-0.891	56.689	-0.817
Dunnit: MTwGWI 11111	0	0.042	73.836	13.669	0.150	0.214	0.385	0.267	57.002	0.451

Table 7: Coherence scores for the Dunnit scenario (rounded).

	OG	OGGen	Sh	ShGen	DM	R	Fi	SZ	SLR	SL
Dunnit: Dunnit<Twin	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE

Table 8: Desideratum satisfaction for the Dunnit scenario.

References

Akiba, K. (2000). Shogenji's probabilistic measure of coherence is incoherent. *Analysis*, *60*(4), 356–359. https://doi.org/10.1093/analys/60.4.356

Bovens, L., & Hartmann, S. (2004). *Bayesian epistemology*. Oxford University Press.

Crupi, V., Tentori, K., & Gonzalez, M. (2007). On Bayesian measures of evidential Support: theoretical and empirical Issues. *Philosophy of Science*, *74*(2), 229–252. https://doi.org/10.1086/520779

Douven, I., & Meijs, W. (2007). Measuring coherence. *Synthese*, *156*(3), 405–425. https://doi.org/10.1007/s11229-006-9131-z

Fenton, N., & Neil, M. (2018). *Risk assessment and decision analysis with bayesian networks*. Chapman; Hall.

Fitelson, B. (2003). A Probabilistic Theory of Coherence. *Analysis*, *63*(3), 194–199.

Glass, D. H. (2002). Coherence, Explanation, and Bayesian Networks. In G. Goos, J. Hartmanis, J. van Leeuwen, M. O'Neill, R. F. E. Sutcliffe, C. Ryan, … N. J. L. Griffith (Eds.), *Artificial Intelligence and Cognitive Science* (Vol. 2464, pp. 177–182). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45750-X_23

Koscholke, J. (2016). Evaluating test cases for probabilistic measures of coherence. *Erkenntnis*, *81*(1), 155–181. https://doi.org/10.1007/s10670-015-9734-1

Meijs, W., & Douven, I. (2007). On the alleged impossibility of coherence. *Synthese*, *157*(3), 347–360. https://doi.org/10.1007/s11229-006-9060-x

Merricks, T. (1995). Warrant entails truth. *Philosophy and Phenomenological Research*, *55*, 841–855.

Olsson, E. J. (2001). Why Coherence Is Not Truth-Conducive. *Analysis*, *61*(3), 236–241.

Roche, W. (2013). Coherence and probability: A probabilistic account of coherence. In M. Araszkiewicz & J. Savelka (Eds.), *Coherence: Insights from philosophy, jurisprudence and artificial intelligence* (pp. 59–91). Dordrecht: Springer.

Schippers, M. (2014). *Probabilistic measures of coherence: From adequacy constraints towards pluralism*. *191*(16), 3821–3845. https://doi.org/10.1007/s11229-014-0501-7

Schippers, M., & Koscholke, J. (2019). A General Framework for Probabilistic Measures of Coherence. *Studia Logica*. https://doi.org/10.1007/s11225-019-09848-3

Scutari, M., & Denis, J.-B. (2015). *Bayesian networks in* ***r***. CRC Press.

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. https://doi.org/10.1093/analys/61.2.147

Shogenji, T. (2006). Why does coherence appear truth-conducive? *Synthese*, *157*(3), 361–372. https://doi.org/10.1007/s11229-006-9062-8

Siebel, M. (2004). On Fitelson's measure of coherence. *Analysis*, *64*, 189–190.

Siebel, M. (2006). Against probabilistic measures of coherence. In *Coherence, truth and testimony* (pp. 43–68). Springer.