# Dahlman and Mackor on Coherence and Probability in Legal Evidence

Olsson, Erik J

*Published in:*
Law, Probability and Risk

*DOI:*
10.1093/lpr/mgz017

2019

*Document Version:*
Peer reviewed version (aka post-print)

Link to publication

*Citation for published version (APA):*
Olsson, E. J. (2019). Dahlman and Mackor on Coherence and Probability in Legal Evidence. *Law, Probability and Risk*, 295–303. https://doi.org/10.1093/lpr/mgz017

*Total number of authors:*
1

**Dahlman and Mackor on Coherence and Probability in Legal Evidence**

Erik J. Olsson

## 1. Introduction

In their very rich and stimulating article, Christian Dahlman and Anne Ruth Mackor (2019) investigate a number of important aspects of coherence in law, outgoing from Van Koppen's scenario theory and Amaya's theory of inference to the most coherence explanation, which in its turn builds on Thagard's analysis of explanatory coherence (e.g. Van Koppen 2011; Amaya 2015; Thagard 2000). A major theme in their article is the extent to which coherence can be reduced to probability. According to the kind of reductionism they investigate, "coherence can be defined in probabilistic terms without loss" (p. 1).[1] However, they are careful to distinguish between a reductive approach to the concept of coherence, on the one hand, and a reductive approach to theories about the use of coherence in actual legal practice – what they call "coherentist theories" – on the other. Their central claim, in this connection, is that "even if coherence is reducible to probability, it does not follow that coherentist theories are reducible to or redundant in relation to probabilistic theories". One aspect of this, which they discuss, is that coherence may have a heuristic role to play in reasoning because people find it easier to reason in terms of coherence, even if, in the end, "a coherence assessment should be in accordance with a probabilistic assessment" (p. 20).

It is impossible in a short commentary to address all the interesting problems and proposals that Dahlman and Mackor's discuss. Instead, I have chosen to focus on two distinct aspects of their work that relate to my own research: their claim that Thagard's principles of coherence can be translated into probabilistic vocabulary and their contention that my own work on coherence exemplifies such a reductive approach to the concept. Regarding the former claim I will investigate the extent to which Dahlman and Mackor have succeeded in their translation, leaving the merits and status of Thagard's rather unconventional, yet highly successful, coherence theory aside (for critical assessments see Olsson 2005, pp. 162-170, and Olsson 2017). My ambition is to offer some constructive criticisms on this part of Dahlman and Mackor's work. As for my own position, I take the opportunity to clarify the (very weak) sense in which I am a reductionist about coherence.

## 2. On the probabilistic translation of Thagard's coherence principles

[1] Unless otherwise indicated page references are to Dahlman and Mackor's article.

An exciting aspect of Dahlman and Mackor's paper is their claim that coherence is reducible to probability.[2] They explain reductionism as follows:

> Reductionism, with regard to coherence and probability, is the view that coherence can be reduced to probability. According to reductionism, everything that is said in terms of coherence in the evaluation of evidence can be said in terms of probability without loss. Advocates of reductionism include Shogenji (1999), Glass (2002), Fitelson (2003) and Olsson (2005).

I will comment on the degree in which I view myself as a reductionist in section 3. Furthermore:

> Reductionism goes back to C. I. Lewis, who proposed that 'congruence' between factual beliefs can be defined in probabilistic terms as positive relevance (Lewis 1946, 338). This general idea has been used by reductionists as a basis for a probabilistic account of coherence (Shogenji 1999, 340; Schubert & Olsson 2013, 35; McGrew 2016, 335). A and B are positive relevant for each other if A increases the probability of B, and B increases the probability of A. So, in this view, A and B are coherent with each other if $P(B \mid A) > P(B)$ and $P(A \mid B) > P(B)$.

In fact, positive relevance is a symmetrical relation, so that $P(B \mid A) > P(B)$ implies $P(A \mid B) > P(B)$. Thus the latter clause is redundant given the former.

Dahlman and Mackor go on to note that probability calculus allows for a clear distinction between two kinds of lack of coherence. We have lack of coherence if two propositions are irrelevant to each other: $P(A \mid B) = P(A)$ (and therefore also $P(B \mid A) = P(B)$). But two propositions can also lack coherence in virtue of being negatively relevant to each other: $P(A \mid B) < P(A)$ (and therefore also $P(B \mid A) < P(B)$). The authors conclude, correctly, that "[t]his is one of several features that makes the probability vocabulary more precise that the coherence vocabulary" (p. 13).

Now, Dahlman and Mackor believe that a theory of coherence broadly in line with Thagard's account and which also incorporates the notion of "inference to the most coherent explanation" has the best chance of capturing coherence in its legally relevant sense. This is the reason why their reductive efforts, in section 3 of their paper, focus on Thagard's coherence principles and how they might be translated into probability theory (Thagard, 2000):[3]

---

[2] Although the two authors are not completely in agreement on this point, as they register in the final conclusion section of their paper. Dahlman asserts that "the coherence vocabulary can be translated into a Bayesian vocabulary be defining coherence as positive relevance" (p. 20), whereas Mackor confesses that she is "still agnostic on this issue" (p. 21, footnote 34).

[3] For the record, Thagard himself believes that probability theory, while being "an immensely valuable tool for making statistical inferences about patterns or frequencies in the world", "is not the appropriate mathematics for understanding human inference in general" (Thagard 2000, pp. 249-250). See also Thagard (2004).

*E1 symmetry*: Explanatory coherence is a symmetrical relation, unlike explanatory relations and relations of conditional probability.

*E2 explanation*:

a) an hypothesis (H) coheres with what it explains: evidence (E) or another hypothesis (H);

b) hypotheses that together explain another proposition (H or E) cohere;

c) the more hypotheses needed to explain something, the lower the degree of coherence.

*E3 analogy*: Similar hypotheses that explain similar pieces of evidence cohere.

*E4 data priority*: Propositions that describe the results of observations (evidence) have a degree of acceptability on their own.

*E5 contradiction*: Contradictory propositions are incoherent with each other.

*E6 competition*: Hypotheses that both explain a proposition but that are not explanatorily connected (cf. principle 2b) are incoherent (i.e. also if they do not contradict each other).

*E7 acceptance*: The acceptability of a proposition in a system depends on its coherence with other propositions.

Dahlman and Mackor note, referring to Amaya (2013), that two additional principles may be needed for the context of criminal laws. First, factual hypotheses that are compatible with innocence have a degree of acceptability of their own. Second, the guilt hypothesis may be accepted only if it is justified to a degree sufficient to satisfy the reasonable doubt standard (pp. 7-8). These two extra principles do not play any role in Dahlman and Mackor's reduction argument, though, and one can argue that this is as it should be since they do not involve the concept of coherence.

Now, principle E1 states that coherence is a symmetrical relation. How can this principle be translated into probabilistic language? Dahlman and Mackor's idea is to define coherence, for two propositions, as positive relevance. Thus, a piece of evidence (E) is coherent with a hypothesis (H) if $P(H \mid E) > P(H)$ and therefore also $P(E \mid H) > P(E)$. Since, as we noted, positive relevance is symmetrical, E1 is satisfied.

So far so good. But already principle E2 a) poses some problems for a reductive approach: "an hypothesis (H) coheres with what it explains: evidence (E) or another hypothesis (H)". What should we mean, in probabilistic terms, by "explains" here? Is the idea that rather we should take explanation as a primitive concept, i.e., a concept that is not further analyzed? What Dahlman and Mackor write, in this connection, is not entirely illuminating:

> In a criminal trial, the main hypothesis is the prosecutor's accusation against the defendant, for example 'the defendant killed the victim by shooting him in the head', and the piece of evidence coherent with this hypothesis could for example be a witness that identifies the defendant as the shooter. This corresponds to Thagard's 'principle of explanation' (E2 sub a) discussed above (section 2.2). The hypothesis 'explains' the evidence.

In what sense this "corresponds to" Thagard's principle is not entirely clear. To put it otherwise: what we need is a probabilistic translation of E2 a) and not merely an example that supposedly "corresponds to" that principle. Later in their paper, in section 4, Dahlman and Mackor write: "[r]eductionism does not claim that explanation can be reduced to probability". If we apply this claim to the present problem, E2 a) should be translated into probability while leaving the notion of explanation itself unanalyzed. If so, a reasonable translation would be simply: "an hypothesis (H) is positively relevant to what it explains: evidence (E) or another hypothesis (H)".

Let us move on to E2 b) – "hypotheses that together explain another proposition (H or E) cohere" – and its probabilistic rendering. Dahlman and Mackor write:

> This is a familiar in criminal evidence, and often referred to as a 'chain'. As an example, H1 could be the hypothesis that the defendant killed the victim, H2 the hypothesis that the defendant was at the crime scene and E a witness that observed the defendant at the crime scene. This corresponds to Thagard's 'principle of explanation' (E2 sub b) discussed above (section 2.2.). Hypothesis that 'together explain' another proposition cohere.

In this case, they explain, "H1, H2 and E are all positively relevant to each other, and are coherent with each other", in conformity with E2 b).

The example is also used as a preliminary in an argument to the effect that E2 c) – "the more hypotheses needed to explain something, the lower the degree of coherence" – can be given a probabilistic interpretation:

> There is always some uncertainty between two hypotheses in a chain, $P(H1 \mid H2) < 1$ and $P(H2 \mid H1) < 1$, otherwise there would be no point in modelling H1 and H2 as separate hypotheses. As Figure 3 shows, the probability of H2 given the evidence is therefore higher than the probability of H1 given the evidence, $P(H2 \mid E) > P(H1 \mid E)$.

This observation leads up to the following conclusion:

> In Figure 3 the chain only contains two hypotheses (H1 and H2), but there are situations where the chain is longer, e.g. H1 – H2 – H3 – H4 – E. In this longer chain, 'explaining E from H1'

needs to move over H2, H3 and H4. Since there is uncertainty between each link in the chain, the probability of H1 given the evidence decreases when the number of hypotheses increases. This corresponds to Thagard's 'principle of simplicity' (E2 sub c) discussed above: 'the more hypotheses needed to explain something, the lower the degree of coherence'.

The problem with this argument is similar to the difficulty we observed in connection with E2 a). We may agree that the fact just mentioned intuitively "corresponds to" Thagard's principle, but it remains unclear what the exact probabilistic translation of that principle should look like. As it stands, E2 c) is formulated in terms of "degree of coherence" and not coherence as matter or all or nothing. Yet, so far only the latter notion of coherence has been given a probabilistic interpretation by Dahlman and Mackor: A and B cohere (in the absolute sense) if and only if $P(B \mid A) > P(B)$.

Can we extrapolate anything about "degree of coherence" from Dahlman and Mackor's example? Taken at face value, their suggestion is that the degree of coherence of a set of hypotheses plus a piece of evidence is essentially related to the probability of the – to use their expression – "main hypothesis" given the evidence, the main hypothesis being in a criminal case the prosecutor's accusation against the defendant. Let us now add to this thought the further proposal that the degree of coherence between A and B can be equated with the average support that one of the propositions confers on the other. If so, Dahlman and Mackor in effect read Thagard's principle as follows: "the more hypotheses needed to explain something, the lower the degree of coherence *between the main hypothesis and the evidence*". This is certainly an interesting proposal, but it is not very close to Thagard's actual formulation of E2 c). Thagard's principle is more naturally read as being about the degree of coherence *of all the propositions involved* and not, or not mainly, about the coherence between two selected propositions.

How does Thagard himself explain the kind of coherence involved in E2? In his book from 2000, Thagard writes (p. 44): "Principle E2, Explanation, describes how coherence arises from explanatory relations: when hypotheses explain a piece of evidence, the *hypotheses cohere with the evidence and with each other*" (my emphasis). Applied to E2 c), this means that in a case in which more hypotheses are need to explain something, this is less so. In order words, it is then to a lesser extent the case that "hypothesis cohere with the evidence and with each other". To show this it does not suffice to show that the degree of coherence or relative relevance between the main hypothesis and the evidence will be lower. It would have to be shown that the degree of coherence or relative relevance between all the propositions involved will generally be lower. One way of making this more precise would be to measure the overall degree of coherence of a set by calculating the *average degree of relative relevance between its elements*. That proposal yields the following probabilistic translation of E2 c):

"the more hypotheses needed to explain something, the lower the average degree or relative relevance between the hypotheses and between the hypotheses and the evidence". Is it true? I don't know. My point here is only that a successful translation of Thagard's coherence principles requires the truth of either this claim or a claim along the same lines.[4]

Yet, Dahlman and Mackor seem to disagree:

> There is a debate among reductionists on how to measure the degree of coherence between two beliefs. Shogenji has proposed that the ratio P(A&B)/P(A)P(B) should be used as a measure (Shogenji 1999, 339). Other ways to measure the degree of coherence have been suggested by Olsson (2002), Fitelson (2003) and Douven & Meijs (2007). The question becomes increasingly complicated when we not only want to measure the degree of coherence between two beliefs, but also the degree of coherence in a set that contains a larger number of beliefs. The various ways to measure the degree of coherence proposed by Shogenji and others have spawned a debate on the merits and shortcomings of the proposed measures (Siebel 2005; Bovens & Hartmann 2005; Moretti & Akiba 2007; Roche 2013; Schippers 2014). Some arguments in this debate are directed at a specific measure. Several participants have argued that there is no objectively superior measure (Bovens & Hartmann 2005; Schippers 2014). This discussion will not be addressed in the present paper.

If what I have concluded above is correct, then Dahlman and Mackor, in their reductive enterprise, need not only a probabilistic measure of the degree of coherence between two beliefs, but also a measure of the degree of coherence in a set that contains a larger number of beliefs. The measure I proposed, very tentatively, above (average degree of relative relevance between each pair of members in the set) at least has the merit of being very broadly in the spirit of Thagard's work.[5]

Thagard's principle E3 states that "similar hypotheses that explain similar pieces of evidence cohere". Dahlman and Mackor are right in objecting that this principle is very vague. In their words: "Unfortunately, Thagard does not explain what 'similar' stands for, and the analogy-feature of Thagard's notion of coherence is therefore too unclear to be translated into probabilistic terms". They observe, moreover, that "Thagard's principles of 'data priority' (E4) and 'acceptance' (E7) do

[4] The proposal made here is related to, but does not seem to coincide with, coherence measures that have been suggested and studied in the literature. It is unusual in its sole focus on the mutual support between pairs of elements of the set, as opposed to, say, the coherence of one element with the rest, along the lines of Lewis (1946). For overviews of coherence measures based on mutual support, see Douven and Mejis (2007), Schippers (2014) and Koscholke (2016).

[5] I write "very broadly" because Thagard himself, although his principle E2 c) clearly relies on a concept of "degree of coherence", has stated that he is skeptical regarding the prospects of actually defining such a concept in a precise fashion within his own formal framework. It is clear that he thinks that his reasons for doubt carry over to probability theory and other formal frameworks. See Thagard (2000, p. 39) for a discussion.

not define coherence" (p. 16, footnote 26), concluding that these principles are not relevant to their concerns.

It remains to consider E5 and E6 from a probabilistic perspective. E5 states, we recall, that "[c]ontradictory propositions are incoherent with each other" and E6 that "[h]ypotheses that both explain a proposition but that are not explanatorily connected … are incoherent". As we noted, there are two kind of incoherence or lack of coherence in probability theory: irrelevance and negative relevance. In the case of contradictory propositions we clearly have incoherence in the sense of negative relevance, so that E5 becomes true on the obviously probabilistic translation. Regarding E6, Dahlman and Mackor write:

> Situation where two (or more) hypotheses together explain the evidence must be distinguished from situations where two hypotheses are negatively relevant for each other, and therefore incoherent. This is the case when H1 and H2 offer competing explanations for the same evidence. Instantiating H1 will decrease the probability of H2, and vice versa. This is known as 'explaining away' in Bayesian networks.

For example, "H1 could be the hypothesis that the defendant shot the victim, H2 the hypothesis that the defendant has been framed by the police, and E a shoe print at the crime scene that matches the defendant's shoe". Dahlman and Mackor conclude, for good reasons, that "[t]his corresponds to Thagard's 'principle of competition (E6) discussed above (section 2.2). Hypotheses that offer competing explanations are incoherent with each other."

**3. Am I a reductionist about coherence and probability?**

As we saw, Dahlman and Mackor view my work on coherence and probability as exemplifying a "reductionist" approach. Drawing this conclusion from the fact that all my later work on coherence, stating with Bovens and Olsson (2000), has been carried out within a probabilistic framework is certainly not unreasonable. Yet, I hesitate to subscribe to Dahlman and Mackor's characterization of reductionism, i.e. to the claim that "everything that is said in terms of coherence in the evaluation of evidence can be said in terms of probability without loss". I will try to explain why.

The methodological framework that I have come to favor over the years, and increasingly committed myself to in writing, is Rudolf Carnap's method of explication. By explication, Carnap meant "the transformation of an inexact, prescientific concept, the explicandum, into a new exact concept, the explicatum" (1950, p. 3). Carnap thought that an explication is successful if the explicatum is exact, fruitful and simple while at the same time being sufficiently similar to the explicandum. A concept is fruitful, according to Carnap, to the extent that it figures in lawlike generalizations, if it is an empirical concept; logical concepts are fruitful in virtue of figuring in formal theorems. An important aspect of Carnap's methodology is that there is, in a sense, no right or wrong when explicating a given concept. Rather, an explication can be more or less useful or satisfactory in a given context or for a given application. A consequence is that an explication that is useful in one context might not be so in another. Thus, the method of explication is, already for that reason, inherently pluralistic (Olsson, 2017; see also Olsson, 2015). But even given one and the same context or application, researchers may come up with different explicate of the explicandum depending on how they weigh the different desiderata. Having said this, Carnap took fruitfulness and exactness to be the most highly valued desiderata, and equally so, followed by similarity and simplicity, in that order.

Now, some definitions of coherence in terms of probability are, in my view, paradigm examples of successful explications. This goes for the definition suggested by Shogenji (1999):

$$C_S(A_1,\ldots,A_n) = \frac{P(A_1 \wedge \ldots \wedge A_n)}{P(A_1) \times \ldots \times P(A_n)}$$

It is also true, I believe, of the measure that was introduced in Olsson (2002) (and independently in Glass 2002):

$$C_O(A_1,\ldots,A_n) = \frac{P(A_1 \wedge \ldots \wedge A_n)}{P(A_1 \vee \ldots \vee A_n)}$$

Both this measures are obviously exact and simple. They have, moreover, been shown to be fruitful in the sense that they figure in many formal theorems relating them to other interesting properties. Thus, Dietrich and Moretti (2005) found that precise conditions under which evidence that supports parts of a theory also supports the theory as a whole can be stated in terms of the "Olsson measure".[6] Angere (2007, 2008) studied the statistical correlation between "higher degree of coherence" and "higher probability" for different measures of coherence using computer simulation, finding that the correlation was particularly high in the case of the Olsson measure. Similarly, Glass (2012) showed by computer simulation that if the Olsson measure is used, then "inference to the most coherent hypothesis" becomes very good at tracking the true hypothesis as well as the most probable hypothesis.[7] Last but not least, it can be formally proved, essentially, that no coherence measures is true conducive in the sense of guaranteeing that a higher degree of coherence is associated with a higher likelihood of truth, not even in a weak *ceteris paribus* sense (Bovens and

[6] The measure is commonly called the "Olsson measure" but some authors refer to it as the "Glass-Olsson measure" or the "Olsson-Glass measure".
[7] Angere's and Glass's results are in line with remarks made by Dahlman and Mackor, in section 5 of their paper, emphasizing the heuristic value of coherence assessments.

Hartmann 2003; Olsson 2005). By implication, this goes for the Thagard inspired measure that I proposed in the previous section, too. That there are no measures that are truth conducive in this sense is a general insight of great epistemological significance, which could not have been established with mathematical precision without a translation of coherence theory into the exact framework of probability.

Whether or not these or other coherence measures as explicata are sufficiently similar to their explicanda is a matter of some dispute and, I believe, depends to some extent on what one takes the relevant explicanda to be. My own take has been to argue that they capture, reasonably well, two pre-systematic senses of coherence: “agreement”, in the case of the Olsson measure, and “striking agreement”, in the case of the Shogenji measure (Olsson, 2002).[8]

The embedding of the concept of coherence in probability theory has overall been a spectacular successful enterprise. This is not to say, however, that probability theory is the only exact framework within with coherence could be explicated. On the face of it, Thagard’s framework offers a different explication of coherence more closely related to the connectionist program in cognitive science. If Dahlman and Mackor are correct, at least the part of Thagard’s work represented by his coherence principles can be given reasonable probabilistic translations. Although I claim to have shown above that the jury has not yet given its final verdict on this matter.

**4. Conclusion**

There is much I can agree with in Dahlman and Mackor’s article. I find their discussion of the scenario approach to coherence enlightening and agree with much of what they say regarding the heuristic value of coherence, to mention just two points of congeniality. My focus here has been on distinct aspects of their paper that relate to my own work: their claim to reduced Thagard’s coherence principles to probability theory, and their categorization of my own work as a similar form of reductionism.

Setting Thagard’s analogy principle aside, Dahlman and Mackor claim to have “demonstrated that Thagard’s notion for coherence in terms of ‘symmetry’, ‘explanation’, ‘simplicity’, ‘contradiction’ and ‘competition’ can be translated into probabilistic vocabulary” (section 3). While I find their project of offering a probabilistic rendering of Thagard’s principles quite intriguing, I believe that there is yet work to be done in this regard. My main concern is principle E2 c), which, on the face of it, requires a concept of “degree of coherence” – indeed, not only for pairs of propositions but for sets of

---

[8] For similar interpretations of coherence measures as explications in Carnap’s sense, see Schubert () and Koscholke (2016).

propositions generally. Thus, I think Dahlman and Mackor are mistaken when they claim the opposite: that they don't need such a concept. On the positive side, I proposed a probabilistic measure of coherence that seems close to Thagard's own explanations of coherence. According to this, very tentative, proposal, the degree of coherence of a set is the average degree of positive relevance between all pairs in the set.

Dahlman and Mackor position me as a "reductionist" regarding coherence, which gave me the opportunity to clarify my position in this regard. It is true that my work has involved translating the notion of coherence, as it is used in epistemology, into probability calculus. However, rather than being a reductionist I view myself as having *explicated* coherence within probability theory, in Rudolf Carnap's sense of "explication". Hence, I believe that translating coherence into probability gives rise to an exact, fruitful and simple concept that is at the same time reasonably close to what we normally mean when we say that propositions "hang well together" or "agree". At the same time, I do not exclude the possibility that there are other ways of making coherence more precise without bringing in probabilities, and that the result of such an alternative endeavor, too, can be a fruitful yet simple account that also respects our most central intuitions. On the contrary, it would hardly be in the spirit of Carnap's pluralist methodology to insist on a particular "reduction" suitable for all purposes and applications.

**References**


Amaya, A. (2015), *The Tapestry of Reason*, Hart, Oxford.

Angere, S. (2007), "The Defeasible Nature of Coherentist Justification," Synthese 157 (3): 321–335.

Angere, S. (2008), "Coherence as a Heuristic," Mind 117 (465): 1–26.

Bovens, L. and Harmann, S. (2003), Bayesian Epistemology, Oxford University Press, Oxford.

Bovens, L., and Olsson, E. J. (2000), "Coherentism, Reliability, and Bayesian Networks", *Mind* 109 (436): 685–719.

Dahlman, C., and Mackor, A. R. (2019), "Coherence and Probability in Legal Evidence", *Law, Probability, and Risk*, this volume.

Dietrich, F., and Moretti, L. (2005), "On Coherent Sets and the Transmission of Confirmation," *Philosophy of Science* 72 (3): 403–424.

Douven, I., and Meijs, W. (2007) "Measuring Coherence", *Synthese*, 2007.

Fitelson, B. (2003), "A Probabilistic Theory of Coherence", *Analysis*.

Glass, D. H. (2002) "Coherence, Explanation and Bayesian Networks", O'Neill, M. et al (eds.), *AICS*.

Koscholke, J. (2016), "Evaluating Test Cases for Probabilistic Measures of Coherence", Erkenntnis 81: 155-181.

Lewis, C. I. (1946), *An Analysis of Knowledge and Valuation*, Open Court, La Salle.

McGrew, L. (2016), "Bayes Factors All the Way – Towards a New View of Coherence and Truth", *Theoria*.

Olsson, E. J. (2002), "What is the Problem with of Coherence and Truth", *Journal of Philosophy*.

Olsson, E. J. (2005) *Against Coherence*, Oxford University Press, Oxford.

Olsson, E. J. (2015), "Gettier and the Method of Explication: A 60 Year Old Solution to a 50 Year Old Problem", Philosophical Studies 172: 57-72.

Olsson, E. J., (2017), "Coherentist Theories of Epistemic Justification", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/spr2017/entries/justep-coherence/

Schippers, M. (2014), "Probabilistic Measures of Coherence", *Synthese*.

Schubert, S. (2012), "Coherence Reasoning and Reliability: A Defense of the Shogenji Measure", Synthese 187(2): 305–319.

Shogenji, T. (1999), "Is Coherence Truth Conducive", *Analysis*.

Thagard, P. (2000), *Coherence in Thought and Action*, MIT Press, Cambridge MA.

Thagard, P. (2004) "Causal Inference in Legal Decision Making – Explanatory Coherence vs. Bayesian Networks" *Applied Artificial Intelligence*, 2004.

Van Koppen, P. J. (2011), *Overtuigend Bewijs,* Nieuw Amsterdam, Amsterdam.