Title
Coherence of Information: What It Is and Why It Matters

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Hartmann, Stephan
Trpin, Borut

Publication Date
2023

Peer reviewed

Coherence of Information: What It Is and Why It Matters

Stephan Hartmann (S.Hartmann@lmu.de)
Munich Center for Mathematical Philosophy, LMU Munich, 80539 Munich (Germany)

Borut Trpin (borut.trpin@lrz.uni-muenchen.de)
Munich Center for Mathematical Philosophy, LMU Munich, 80539 Munich (Germany)

Abstract

Coherence considerations play an important role in science and in everyday reasoning. However, it is unclear what exactly is meant by coherence of information and why we prefer more coherent information over less coherent information. To answer these questions, we first explore how to explicate the dazzling notion of "coherence" and how to measure the coherence of an information set. To do so, we critique prima facie plausible proposals that incorporate normative principles such as "Agreement" or "Dependence" and then argue that the coherence of an information set is best understood as an indicator of the truth of the set under certain conditions. Using computer simulations, we then show that a new probabilistic measure of coherence that combines aspects of the two principles above, but without strictly satisfying either principle, performs particularly well in this regard.

Keywords: Reasoning and Argumentation; Coherence; Truth; Formal Epistemology

Introduction

Highly coherent information sets seem more plausible than information sets whose elements do not hang together well. As an example consider the following hypothetical scenario: Three studies have been conducted on the effects of caffeine on human health. Study 1 reports that caffeine consumption is associated with improved cognitive performance, Study 2 reports that caffeine consumption is associated with increased anxiety levels, and Study 3 reports that caffeine consumption is associated with disrupted sleep patterns. In another scenario, Study 1 also reports improved cognitive performance, Study 2 reports reduced levels of anxiety, and Study 3 reports no significant effect on sleep patterns. The three studies mentioned in the first scenario are in tension with each other as, given our background knowledge, anxiety and disrupted sleep negatively affect cognitive performance. On the other hand, the three studies are coherent in the second scenario. The degree of coherence of the three studies in question does not mean that they are necessarily true, but yet the higher coherence makes the studies in the second scenario seem more plausible overall than the studies in the first scenario.

It is very plausible that coherence considerations like those play an important role in science and everyday reasoning (see, e.g., Harris & Hahn, 2009; Hahn, Harris, & Corner, 2016) in a descriptive sense. However, it is not clear whether they also have (or should have) any normative epistemological significance. Is it reasonable to take the coherence of an information set to be an indicator of how much we should believe the information set in question? Note that we follow Bovens and Hartmann (2003, pp. 10-11) in taking coherence to be a property of information *sets* and not, e.g., of propositions. Formally, if we obtain the information items $\mathrm{R}_1, \ldots, \mathrm{R}_\mathrm{n}$ from n independent and partially reliable sources, then $\mathbf{S} = \{\mathrm{R}_1, \ldots, \mathrm{R}_\mathrm{n}\}$ is an information set over which a (subjective) probability distribution is defined.

To answer questions related to the normative role of coherence considerations, we also need to look for a way to measure the coherence of an information set. The literature in formal epistemology provides a number of probabilistic measures of coherence that are supposed to do just this (e.g., Shogenji, 1999; Glass, 2002; Olsson, 2002; Fitelson, 2003; Bovens & Hartmann, 2003; Douven & Meijs, 2007; Schupbach, 2011; Koscholke, Schippers, & Stegmann, 2019). But the measures differ in their assessments. How should we then determine which measure of coherence (if any) is most fit for determining the normative role of coherence considerations?

We find three types of arguments in the literature. First, proposed measures are confronted with test cases for which we have a clear intuition (see Koscholke, 2016). Unfortunately, these test cases usually involve only information pairs and triples as it is difficult to develop reliable intuitions for larger information sets. Second, empirical studies are conducted to determine which coherence measure best represents our coherence intuitions (see, e.g., Harris & Hahn, 2009; Koscholke & Jekel, 2017). In addition to the controversial is-to-ought inference, the results obtained also cannot be used as a normative guide because they are too diverse. Third, we may refer to plausible normative principles that should be satisfied by the proposed coherence measures. Interestingly, it turned out that the two most important normative principles, **Agreement** and **Dependence**, are mutually exclusive (Schippers, 2014). We will examine these two principles in more detail below and argue that they are too strict for larger information sets. Finally, we will argue that the best way to evaluate a proposed coherence measure is to show that it serves a desirable function, namely that of helping us figure out which information sets are true and which are false.

Two Main Normative Principles

Not all information sets are equally coherent. Consider two information sets, $\mathbf{S}$ and $\mathbf{S}'$. $\mathbf{S}$ includes the following information: "The weather forecast predicts heavy rain tonight", "The

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

streets in this neighborhood tend to flood during heavy rain", and "The drainage system in this neighborhood is outdated". $\mathbf{S}'$, on the other hand, includes: "The weather forecast predicts heavy rain tonight", "The streets in this neighborhood tend to flood during heavy rain", and "This neighborhood recently completed a major overhaul of its drainage system".

Both information sets seem coherent, but the information in $\mathbf{S}$ fits together more strongly than in $\mathbf{S}'$. In $\mathbf{S}$, the prediction of heavy rain, the tendency for the streets to flood, and the outdated drainage system all support the conclusion that the streets are likely to flood tonight. In $\mathbf{S}'$, the information about the updated drainage system weakens that conclusion.

Now suppose we add "The local river has recently undergone dredging and widening to prevent flooding" to $\mathbf{S}$ and "The recent overhaul of the drainage system was not completed properly and may malfunction" to $\mathbf{S}'$. It is then not clear which of the two information sets is more coherent. Some authors even claim that it is not always possible to say which set is more coherent than another (see, e.g., Bovens & Hartmann, 2003). This may perhaps also be the case here.

Instead of referring to intuitions about specific test cases, an alternative approach to determining which measure of coherence is the most adequate, relies on normative principles that have a certain intuitive appeal. The idea is that any adequate measure of coherence should satisfy such principles. For instance, consider the so-called Principle of Agreement (hereafter simply **Agreement**). The principle goes back to Bovens and Olsson (2000) and has been revived in, e.g., Schippers (2014) and Koscholke et al. (2019). It roughly states that increasing the conditional probabilities of all information items, given other information items, should increase the coherence of the information set because there is then more mutual support.

Another principle that has an intuitive appeal is the Principle of Dependence (hereafter **Dependence**). According to Brössel (2015), the principle might be tracked all the way back to Keynes (1921), although its application to coherence has been particularly clear since the (re)introductions by Shogenji (1999) and Fitelson (2003). **Dependence** states, in simple terms, that an information set is absolutely coherent (incoherent) if the information items are positively (negatively) correlated.

Before we give formal definitions of the two principles, a key concept needs to be defined:

Definition 1. *A probability distribution P is defined over a set of propositional variables $V := \{H_1, \dots, H_n\}$ with the values* H_i *and* $\neg\mathrm{H}_i$ *for all $i = 1, \dots n$.*

(i) *V is* independent (relative to *P*) *iff* $P\big(\bigwedge_{i\in I} \mathrm{H}_i\big) = \prod_{i\in I} P(\mathrm{H}_i)$ *for all non-empty subsets $I \subseteq \{1, \dots, n\}$.*

(ii) *V is* positively correlated (relative to *P*) *iff* $P\big(\bigwedge_{i\in I} \mathrm{H}_i\big) \geq \prod_{i\in I} P(\mathrm{H}_i)$ *for all non-empty subsets $I \subseteq \{1, \dots, n\}$ and at least one of the "$\geq$" is a "$>$".*

(iii) *V is* negatively correlated (relative to *P*) *iff* $P\big(\bigwedge_{i\in I} \mathrm{H}_i\big) \leq \prod_{i\in I} P(\mathrm{H}_i)$ *for all non-empty subsets $I \subseteq \{1, \dots, n\}$ and at least one of the "$\leq$" is a "$<$".*

The two principles can then be described in the following way (roughly following Koscholke et al., 2019):

Definition 2. ***(Dependence).*** *Given a measure of coherence Coh, we say that it satisfies Dependence if there is a threshold τ such that for any information set $\mathbf{S}$:*

- $Coh(\mathbf{S}) > \tau$ *if $\mathbf{S}$ is positively correlated.*
- $Coh(\mathbf{S}) = \tau$ *if $\mathbf{S}$ is independent.*
- $Coh(\mathbf{S}) < \tau$ *if $\mathbf{S}$ is negatively correlated.*

Definition 3. ***(Agreement).*** *Let us assume that the following inequality holds for all non-empty disjoint subsets of conjunction $\mathbf{S}', \mathbf{S}''$ of some information set $\mathbf{S}$ for two probability distributions P_1 and P_2:*

$$P_1\left(\bigwedge_{s_j\in\mathbf{S}'} s_j \mid \bigwedge_{s_m\in\mathbf{S}''} s_m\right) > P_2\left(\bigwedge_{s_j\in\mathbf{S}'} s_j \mid \bigwedge_{s_m\in\mathbf{S}''} s_m\right).$$

Given a coherence measure Coh, we say that it satisfies Agreement if it then also holds that: $Coh_{P_1}(\mathbf{S}) > Coh_{P_2}(\mathbf{S})$ for the measurements of coherence by Coh relative to probability distributions P_1 and P_2, respectively.

Interestingly, as proven by Schippers (2014) and further developed by Koscholke et al. (2019), **Agreement** and **Dependence** mutually exclude each other – any measure that satisfies one cannot satisfy the other principle. This is bad news because both principles have an intuitive appeal. In response, Koscholke et al. (2019) argues for pluralism and suggests determining in which contexts which of the two principles makes sense and then finding the best measure that satisfies the principle in question. We, on the other hand, believe that both principles are generally untenable. For this purpose, let us consider the following counterexample to **Agreement**:

NEW PRODUCT There is a company that produces a popular product, and Mr A is in charge of its marketing. Ms B is responsible for product design. Consider two versions:

1. Mr A and Ms B tend to have different opinions on how to approach the market. Mr A is very likely to be involved in the decision-making process. If Mr A is involved, Ms B may reconsider her design. If he is not involved, however, Ms B is very likely to go ahead with her ideas. Suppose that the probability distribution is defined by the following three values: $P(\mathrm{A}) = P(\mathrm{B}|\neg\mathrm{A}) = .9$ and $P(\mathrm{B}|\mathrm{A}) = .6$, where A stands for Mr A being involved in the decision-making process and B for Ms B going ahead with her design ideas.
2. Ms B is very unlikely to go ahead with her design ideas if Mr A is not involved, and more likely than not to go on if he is involved. However, Mr A has been on a sick leave, so it is unlikely that he is involved in the upcoming decision-making process.

Suppose the probability distribution is defined by the following values: $P(\mathrm{A}) = P(\mathrm{B}|\neg\mathrm{A}) = .1$ and $P(\mathrm{B}|\mathrm{A}) = .55$.

In both cases, two independent witnesses each give the reports:

R_1: Mr A was involved in the decision-making process.

R_2: Ms B went ahead with her design ideas.

In which of the two situation do the two reports *fit together* better?

The question is not in which case the two reports are more likely, but rather in which situation the two reports fit together better or, equivalently, in which situation there is less tension between them. Taking this into account, it seems that the two reports R_1 and R_2 fit together better in situation 2. In situation 1, both Mr A and Ms B are likely to work on the new product, but A's involvement dissuades B from going on with her ideas. In situation 2, on the other hand, A's involvement encourages B's, and as **Dependence** requires, the reports are therefore more coherent.

However, any measure that respects **Agreement** will give an opposite response because all the mutual conditional probabilities are greater in situation 1 than in situation 2: $P_1(\mathrm{A}|\mathrm{B}) = .86 > P_2(\mathrm{A}|\mathrm{B}) = .38$ and $P_1(\mathrm{B}|\mathrm{A}) = .6 > P_2(\mathrm{B}|\mathrm{A}) = .55$. So, according to these measures, the first situation is the more coherent one. To us this seems wrong because in the first situation there is more tension between A and B. Hence, **Agreement** and any measure of coherence that satisfies it should be rejected.

But what about the principle of **Dependence**? Since the principles of **Agreement** and **Dependence** are known to be mutually exclusive, our counterexample offers the possibility of motivating coherence measures that satisfy **Dependence**. However, it turns out that **Dependence** is too strict for larger information sets, and thus should not be a general desideratum for an acceptable coherence measure. This is essentially because there are information sets for which all propositions are positively correlated but have almost no overlap. Without a sufficient amount of overlap, however, there can be no "hanging together" in the first place, and thus no coherence.

We can show the issue with the following example. Suppose that there is a town where it rains frequently. Let R represent that it is raining in the town, let B represent that a person in the town is reading a book, and let C represent that a person in the town is wearing a raincoat. Suppose that the probability of R (rain) is 0.5. The probability of B (reading a book) given R is 0.2, meaning that when it rains, 20% of people are expected to read books. The probability of B given not-R (no rain) is 0.1 because fewer people read books on a non-rainy day. The probability of C (wearing a raincoat) given R is 0.9 and only 0.3 given not-R. Assuming that the propositional variable R probabilistically screens off the propositional variable B from the propositional variable C (or, in causal language, C is the common cause of B and C), we have enough information to calculate the joint probability distribution over all three variables.

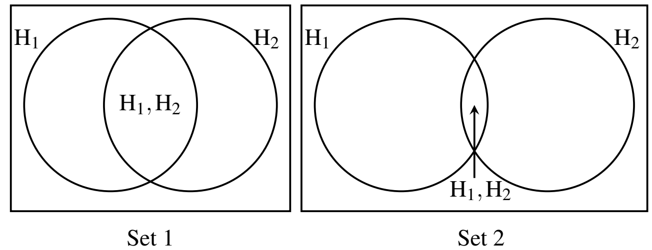

Figure 1: The relative overlap of the two information items $\mathrm{H}_1, \mathrm{H}_2$ in information set 1 is larger than in set 2.

Note that the set $\mathbf{S} = \{\mathrm{R}, \mathrm{B}, \mathrm{C}\}$ is positively dependent, although there is very little relative overlap of the three information items because most people of the town don't read books regardless of the weather or what they are wearing. Hence, it is not clear whether the set $\mathbf{S}$ should be considered as absolutely coherent: the set $\mathbf{S}' = \{\mathrm{R}, \neg\mathrm{B}, \mathrm{C}\}$ (rain, not reading a book, but wearing a raincoat) has, after all, a much higher degree of relative overlap.

Finally, even if one does not agree with our verdict regarding **Dependence** and **Agreement**, it should be noted that the two principles are practically inapplicable when we consider the coherence of larger information sets. The principles only provide guidance for cases where all non-empty subsets of an information set are correlated (**Dependence**) in the same way or for cases where the conditional probabilities of information conjuncts from specifically defined subsets are all greater under one probability function than under another (**Agreement**). For the vast majority of sets, these conditions do not hold as they are very demanding. Therefore, these principles are void for the majority of information sets.

For instance, if we use a few lines of code to generate 10,000 random probability distributions for various sizes of information sets, we find the following: For sets with three information items, **Dependence** only applies in 2400 cases. For sets with four information items, the number of applicable cases drops to 227, and further to nine for sets with five, and all the way down to zero for sets with six information items. This clearly shows that the larger a random information set is, the more likely it is that it is neither independent nor positively/negatively correlated as defined by the conditions of Definition 5. Notably, any information pair is either independent or (positively/negatively) correlated. This provides a case in point of using **Dependence** as a normative principle of coherence for information pairs. Note also that **Agreement** involves a comparison of two probability distributions, so we assume that it provides normative guidance in even fewer cases than **Dependence**. In summary, the two principles are unlikely to be of much use in practice.

Coherence as the Degree of Relative Overlap

To motivate our own proposal of how and why to measure coherence, we first consider the class of relative overlap measures of coherence. These measures build on the intuitions that (i) an information set consisting of two inconsistent information items (whose joint probability is therefore zero), is minimally coherent and that (ii) an information set consisting of two information items perfectly overlapping in probability space is maximally coherent. Accordingly, coherence measures the relative overlap of propositions in probability space, as illustrated by the Figure 1). This leads to the following measure of the coherence of an information set $\mathbf{S} := \{\mathrm{H}_1, \ldots, \mathrm{H_n}\}$ relative to a probability distribution P (Glass, 2002; Olsson, 2002).[1]

$$\mathrm{Coh_{OG}}(\mathbf{S}) := \frac{P(\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_n)}{P(\mathrm{H}_1 \vee \mathrm{H}_2 \vee \cdots \vee \mathrm{H}_n)}. \tag{1}$$

Note that the joint probability (numerator) cannot increase and the probability of a disjunction (denominator) cannot decrease when new information is added. Therefore, new information cannot increase the coherence as measured by $\mathrm{Coh_{OG}}$. The following counterexample shows that this is clearly wrong. Consider an information set $\mathbf{S}$ consisting of LS: "John loves to eat steak" and V: "John is a vegetarian". Obviously, $\mathbf{S}$ is not very coherent since vegetarians are not particularly keen on steaks. An overlap measure such as $\mathrm{Coh_{OG}}$ captures this correctly. However, adding another information item H: "John became a vegetarian a week ago for health reasons", should increase the coherence since H reduces the tension between LS and V, so the information fits together very well. Unfortunately, however, $\mathrm{Coh_{OG}}$ comes to a different conclusion for the aforementioned reason.

According to Meijs (2005, 2006), we should instead consider how much relative overlap there is among all non-empty non-singleton subsets of some set $\mathbf{S}$ and then take the average value of coherence obtained in this way. This idea gives rise to the measure $\mathrm{Coh_{OG'}}$,

$$\mathrm{Coh_{OG'}}(\mathbf{S}) = \frac{1}{m}\sum_{i=1}^{m} \mathrm{Coh_{OG}}(\mathbf{S_i'}), \tag{2}$$

where $\mathbf{S_i'}$ represents any of the m non-empty non-singleton subsets of the information set $\mathbf{S}$. This measure resolves the above-mentioned counterexample. However, it cannot judge a given information set $\mathbf{S}$ as more coherent than its most coherent two-element subset (Koscholke & Schippers, 2016).

To overcome this problem, Koscholke et al. (2019) propose a new measure, which combines relative overlap and average mutual support intuitions (Douven & Meijs, 2007). The idea is that to assess how coherent an information set $\mathbf{S}$ is, we need to consider the average relative overlap of non-empty disjoint subsets of conjunctions in $\mathbf{S}$. Formally:

$$\mathrm{Coh_{OG^*}}(\mathbf{S}) = \frac{1}{k}\sum_{i=1}^{k} \mathrm{Coh_{OG}}\left(\bigwedge_{s_j \in \mathbf{S}'} s_j, \bigwedge_{s_m \in \mathbf{S}''} s_m\right)_{\mathrm{i}} \tag{3}$$

where $\mathbf{S}'$ and $\mathbf{S}''$ are subsets as described above.

Unfortunately, the measure $\mathrm{Coh_{OG^*}}$ also falls short. Besides the fact that it satisfies **Agreement**, which we take to not be plausible as the NEW PRODUCT example from the previous section shows, it is also practically untenable. The measure is defined as the average of conjunctions from all respective subsets. Hence, for n information items, the measure averages over $[(3^n - 2^{n+1}) + 1]/2$ calculations of relative overlaps (each calculated by $\mathrm{Coh_{OG}}$) in specifically defined subsets (Koscholke et al., 2019). This means that the computational load exponentially increases when the set under consideration increases and the measure is therefore computationally intractable and hence cognitively implausible.

Combining Relative Overlap and Probabilistic Relevance

To construct a more promising measure based on relative overlap, we first consider the simplest independence deviation measure of coherence (Shogenji, 1999):

$$\mathrm{Coh_{Sh}}(\mathbf{S}) := \frac{P(\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_n)}{P(\mathrm{H}_1)P(\mathrm{H}_2)\cdots P(\mathrm{H}_n)} \tag{4}$$

This measure suffers from its own problems (for some standard counterexamples see, e.g., Fitelson, 2003; Bovens & Hartmann, 2003). However, it is easy to see that it satisfies **Dependence**. If the information set under consideration is positively (negatively) correlated, then the numerator is greater (lesser) than the denominator and the measure will judge an information set to have coherence above (below) the threshold value of 1. If it is independent, then the numerator and the denominator are equal and the coherence is exactly 1.

To continue, note that $\mathrm{Coh_{Sh}}(\mathbf{S})$ is defined as the ratio of the joint probability of the information in $\mathbf{S}$ and the probability of the same information assuming they were probabilistically independent and had the same marginal probabilities. Let us now introduce a new concept that will prove useful as we proceed.

Definition 4. *A probability distribution P is defined over a set of propositional variables $V := \{H_1, \ldots, H_n\}$. The associated probability distribution $\tilde{P}$ satisfies the following conditions: (i) $\tilde{P}$ is defined over the same set V; (ii) V is independent relative to $\tilde{P}$; (iii) $\tilde{P}(\mathrm{H}_i) = P(\mathrm{H}_i)$ for all $i = 1, \ldots, n$.*

With this definition, the Shogenji measure can then be written as $\mathrm{Coh_{Sh}}(\mathbf{S}) = P(\mathbf{S})/\tilde{P}(\mathbf{S})$. It therefore follows from starting with one of the simplest (but not very convincing, see Olsson, 2021) prima facie measures of coherence–viz. the joint probability of $\mathbf{S}$, i.e. $\mathrm{coh}_P^{(0)}(\mathbf{S}) := P(\mathbf{S})$–and then normalizing it by $\mathrm{coh}_{\tilde{P}}^{(0)}(\mathbf{S}) = \tilde{P}(\mathbf{S})$. Because we use the associated probability distribution $\tilde{P}$ in this expression, we can then say that $\mathrm{Coh_{Sh}}(\mathbf{S})$ is the measure of coherence associated with the prima facie measure $\mathrm{coh}_P^{(0)}(\mathbf{S})$. Generalizing from this example, the following definition specifies how one can construct an improved measure of coherence, which takes independence deviation intuitions into account:

[1] When appropriate, we use the convention of representing the conjunction $\mathrm{H}_1 \wedge \mathrm{H}_2 \wedge \cdots \wedge \mathrm{H_n}$ as $\mathrm{H}_1, \ldots, \mathrm{H}_n$. Note also that in what follows we no longer explicitly mention that the coherence of an information set is relative to a probability distribution P.

Definition 5. *Let* $\mathbf{S}$ *be an information set and P be a probability distribution defined over the corresponding set of propositional variables. Furthermore, let* $\tilde{P}$ *be the associated probability measure and let* coh_P *be a prima facie measure of coherence (relative to P). Then*

$$\mathrm{Coh}_P(\mathbf{S}) := \frac{\mathrm{coh}_P(\mathbf{S})}{\mathrm{coh}_{\tilde{P}}(\mathbf{S})} \tag{5}$$

is the associated measure of coherence if $\mathrm{coh}_{\tilde{P}}(\mathbf{S}) > 0$.

Let us now construct the associated measure of coherence from the Olsson–Glass measure $\mathrm{Coh}_{\mathrm{OG}}$. The resulting measure provides a promising compromise between probabilistic relevance and relative overlap. We obtain:

$$\mathrm{Coh}_{\mathrm{OG}^+}(\mathbf{S}) \;:=\; \frac{P(\mathrm{H}_1,\ldots,\mathrm{H}_n)}{P(\mathrm{H}_1\vee\cdots\vee\mathrm{H}_n)}\bigg/\frac{\tilde{P}(\mathrm{H}_1,\ldots,\mathrm{H}_n)}{\tilde{P}(\mathrm{H}_1\vee\cdots\vee\mathrm{H}_n)}. \tag{6}$$

This measure identifies the value of coherence of an information set with the ratio of the relative overlap and the relative overlap that would obtain if the information set were independent. In contrast to other relative overlap measures considered so far, we can show that $\mathrm{Coh}_{\mathrm{OG}^+}$ satisfies **Dependence** for information pairs and triples.

Proposition 1. *An agent considers information items* H_1, H_2, *and* H_3 *with a prior probability distribution P defined over the corresponding propositional variables. Let* $\mathbf{S}_2 := \{\mathrm{H}_1,\mathrm{H}_2\}$ *and* $\mathbf{S}_3 := \{\mathrm{H}_1,\mathrm{H}_2,\mathrm{H}_3\}$. *Then the following hold for* $\mathbf{S} = \mathbf{S}_i$ *with* $i = 2, 3$*: (i)* $\mathrm{Coh}_{\mathrm{OG}^+}(\mathbf{S}) > 1$ *if* $\mathbf{S}$ *is positively correlated; (ii)* $\mathrm{Coh}_{\mathrm{OG}^+}(\mathbf{S}) = 1$ *if* $\mathbf{S}$ *is independent; (iii)* $\mathrm{Coh}_{\mathrm{OG}^+}(\mathbf{S}) < 1$ *if* $\mathbf{S}$ *is negatively correlated.*

This is a significant result because it shows that a relative overlap measure is able to take probabilistic relevance considerations into account, although our base measure of coherence $\mathrm{Coh}_{\mathrm{OG}}$ does not satisfy **Dependence** (Schippers, 2014, p. 3840). As we have just seen, however, $\mathrm{Coh}_{\mathrm{OG}^+}$ satisfies it even for information triples. However, it turns out that **Dependence** does not hold in general for $\mathrm{Coh}_{\mathrm{OG}^+}$. We believe that this is a welcome result because **Dependence** is a very strict principle and it is not clear whether it is reasonable to request it for information sets of any size. However, it is very plausible for smaller information sets. We therefore conclude that our new coherence measure balances well both relative overlap and dependence (or relevance) considerations without strictly satisfying the underlying principles.

Coherence as an Indication of Truth

While principles like **Agreement** and **Dependence** have some plausibility for small information sets, it is not clear how to evaluate coherence measures for larger sets. The larger sets are also problematic for subset-based measures such as $\mathrm{Coh}_{\mathrm{OG}^*}$, since they become computationally practically intractable when the size of the information sets considered is increased. Similarly, test cases are not suitable for evaluating coherence measures in general because they involve only small information sets (see Koscholke, 2016). Therefore, we propose to evaluate coherence measures for larger n in terms of how well they satisfy a particular *function*. Specifically, we propose to evaluate coherence measures by how well they help us identify true information sets. This claim does not seem to fit with the extensive literature on the truth-conduciveness of coherence and the impossibility results proved therein, according to which coherence cannot be an indicator of truth, at least not without further ceteris paribus conditions (for a review, see Olsson, 2021, Sections 7 and 8). And yet it is useful to examine the truth-tracking properties of coherence measures. In doing so, we will find that some measures perform better than others.

Douven (2021) recently demonstrated that we can consider how well various probabilistic measures of confirmation discriminate between true and false hypotheses by means of computer simulations. In answering the puzzle regarding the truth-tracking abilities of coherence, we can follow this approach and adapt it for our present needs. We also note the previous simulations-based research on coherence and truth-tracking, in particular Angere (2007, 2008) and Glass (2012). However, our interest here is not whether higher coherence of an information set implies higher probability on average. Instead, we focus on how well the coherence of an information set distinguishes true and false information sets. Accordingly, Douven's (2021) method is more appropriate here.

To proceed, we note that all relative overlap measures provide a value which describes how coherent a given information set $\mathbf{S}$ is. The information set can also be described in binary terms as true if all information items are true, and false otherwise. We can then use statistical techniques to find how well different measures of coherence are able to discriminate true and false information sets. This provides insight into which measures of coherence are better indicators of truth.

The procedure of our simulations may be roughly described as follows: We generate *n* possible worlds over which we define a random probability distribution and select one of the worlds to be true (the actual world). We then randomly generate the information set $\mathbf{S} = \{\mathrm{H}_1,\ldots,\mathrm{H}_\mathrm{n}\}$, which may or may not be true (if the actual world is in the subset of all information), and the set of true evidence $\mathbf{E} = \{\mathrm{E}_1,\ldots,\mathrm{E}_\mathrm{n}\}$, all of which include the actual world. Finally, to establish a reasonable connection to the truth, we require that each information is positively correlated with true evidence, i.e., that $P(\mathrm{H_i}|\mathrm{E_i}) > P(\mathrm{H_i}|\neg\mathrm{E_i})$ for all $i = 1,\ldots,n$.

After generating the information set $\mathbf{S}$ and a probability distribution as described, we can calculate how coherent $\mathbf{S}$ is according to various measures of coherence, and after 100 repetitions estimate which measure of coherence provides the best model for discriminating the truth and falsehood of the information set by means of area under curve (AUC).

The value of AUC is usually understood as giving us the probability that some independent continuous measure discriminates dependent binary categories (here: true/false). To simplify, an AUC score of 1 suggests that all true information sets were measured as highly coherent and vice versa for

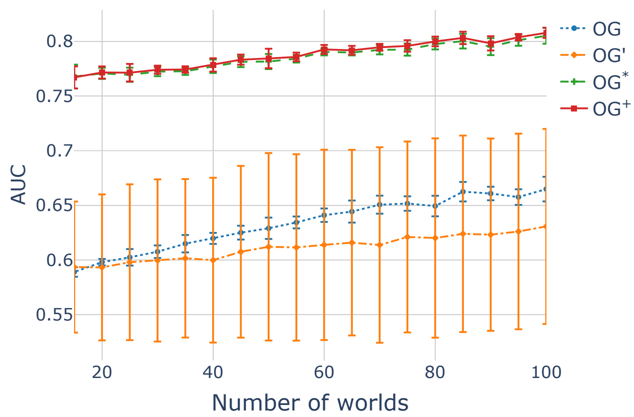

Figure 2: Average AUC values for different measures of coherence, averaged over all simulations with seven information items. The error bars correspond to standard deviations.

false information sets (incoherent). A score of 0.5 suggests that a measure is about as reliable in predicting whether an information set is true or false as a random guess. Finally, for each of the parameters (prior probability of the information set between .1 and .9 in .2 increments and varying cardinality of the set), we repeat the simulations 100 times.

We leave a more thorough simulation-based investigation for our future research. However, for the present needs, it is sufficient to show what happens when we simulate information sets with seven information items. We avoid larger information sets because our results are relatively stable when we increase the simulated information sets, and because the computations become increasingly demanding for measures $\mathrm{Coh}_{\mathrm{OG}'}$ and $\mathrm{Coh}_{\mathrm{OG}^*}$. Note that the differences among the measures decrease if the size of the information set decreases.

Figure 2 shows the average AUC values over 100 simulations with seven information items over all above-described prior probabilities of the sets. It is immediately clear from the figure that our new measure $\mathrm{Coh}_{\mathrm{OG}^+}$ comes out as the top contender with respect to its truth-tracking abilities. Admittedly, it is not a significantly stronger indicator of truth compared to $\mathrm{Coh}_{\mathrm{OG}^*}$ (Koscholke et al., 2019), but note that with seven information items, our measure only requires two calculations ($\mathrm{Coh}_{\mathrm{OG}}$ for P and for $\tilde{P}$), while $\mathrm{Coh}_{\mathrm{OG}^*}$ needs to determine the calculation of 966 subsets. As noted, we get similar results also for sets with five and six information items. For smaller sets, our new measure also comes out as the best on average, but it is not significantly better than the rest. Crucially, this suggests that our measure is a reliable indicator of truth despite its computational simplicity. The relative success of the other measures, however, provides a more general point in favor of coherentist epistemology.

We can conclude that combining an overlap measure with probabilistic relevance considerations not only satisfies desirable properties, but also increases the likelihood that we end up believing true information under certain plausible assumptions (e.g., that the information under consideration is positively correlated with the true evidence).[2]

Conclusion

Our initial question was why we prefer more coherent information to less coherent information. To address this question, we first tried to determine what is meant by "coherence." To make precise the rather vague intuition that coherent information fits together well, we first considered two principles that promise to facilitate the measurement of the coherence of an information set. Unfortunately, it turned out that these prima facie plausible principles are not compatible with each other and individually, at least for large amounts of information, not very plausible. We have therefore proposed to assess the coherence of an information set according to how it helps us reason and deliberate. In particular, we have proposed to judge a coherence measure by how much it helps us identify true information sets. Our computer simulations then showed that the new coherence measure $\mathrm{Coh}_{\mathrm{OG}^+}$, which accounts for both of the above principles in some way (without accounting for them exactly), performs particularly well in this regard despite its computational simplicity.

Appendix

Proof of Proposition 1

Let's begin with $\mathbf{S_2} = \{\mathrm{H}_1, \mathrm{H}_2\}$ and introduce the following shorthands: $\alpha_1 := P(\mathrm{H}_1) + P(\mathrm{H}_2)$, $\beta_1 := \alpha_1$, $\alpha_2 := P(\mathrm{H}_1, \mathrm{H}_2)$ and $\beta_2 := P(\mathrm{H}_1)P(\mathrm{H}_2)$. As $\mathbf{S_2}$ is positively correlated, $\alpha_2 > \beta_2$. We proceed under the assumption that all probabilities are in $(0,1)$. Then we obtain

$$\begin{aligned}
\mathrm{Coh}_{\mathrm{OG}^+}(\mathbf{S_2}) > 1 \quad &\Leftrightarrow \quad \frac{\alpha_2}{\alpha_1 - \alpha_2} > \frac{\beta_2}{\beta_1 - \beta_2} = \frac{\beta_2}{\alpha_1 - \beta_2} \\
&\Leftrightarrow \quad \alpha_1 \alpha_2 > \alpha_1 \beta_2 \\
&\Leftrightarrow \quad \alpha_2 > \beta_2,
\end{aligned}$$

which holds by assumption.

The proof for $\mathbf{S_3} = \{\mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3\}$ proceeds accordingly. Let's first define $\alpha_1 := P(\mathrm{H}_1) + P(\mathrm{H}_2) + P(\mathrm{H}_3)$, $\beta_1 := \alpha_1$, $\alpha_2 := P(\mathrm{H}_1, \mathrm{H}_2) + P(\mathrm{H}_1, \mathrm{H}_3) + P(\mathrm{H}_2, \mathrm{H}_3)$, $\beta_2 := P(\mathrm{H}_1)P(\mathrm{H}_2) + P(\mathrm{H}_1)P(\mathrm{H}_3) + P(\mathrm{H}_2)P(\mathrm{H}_3)$, $\alpha_3 := P(\mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3)$ and $\beta_3 := P(\mathrm{H}_1)P(\mathrm{H}_2)P(\mathrm{H}_3)$. Note that $\alpha_2 > \beta_2$ and $\alpha_3 > \beta_3$. Then we obtain

$$\begin{aligned}
\mathrm{Coh}_{\mathrm{OG}^+}(\mathbf{S_3}) > 1 \quad &\Leftrightarrow \quad \frac{\alpha_3}{\alpha_1 - \alpha_2 + \alpha_3} > \frac{\beta_3}{\alpha_1 - \beta_2 + \beta_3} \\
&\Leftrightarrow \quad \alpha_3(\alpha_1 - \beta_2) > \beta_3(\alpha_1 - \alpha_2).
\end{aligned}$$

This holds because (i) $\alpha_3 > \beta_3$ and (ii) $\alpha_1 - \beta_2 > \alpha_1 - \alpha_2$ is equivalent to $\alpha_2 > \beta_2$, both of which hold by assumption.

The corresponding proofs for negatively correlated and for independent information sets obtain by following the same steps but with "=" (for independent) and "<" (for negatively correlated) information sets instead of ">". □

[2]The code used to conduct the simulations as well as additional plots and data are available here: https://github.com/philosophy-simul/truth-tracking-coherence.

References

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., & Hartmann, S. (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.

Bovens, L., & Olsson, E. J. (2000). Coherentism, reliability and Bayesian networks. *Mind*, *109*(436), 685–719.

Brössel, P. (2015). Keynes's coefficient of dependence revisited. *Erkenntnis*, *80*(3), 521–553.

Douven, I. (2021). Tracking confirmation. *Philosophy of Science*, *88*(3), 398-414.

Douven, I., & Meijs, W. (2007). Measuring coherence. *Synthese*, *156*(3), 405–425.

Fitelson, B. (2003). A probabilistic theory of coherence. *Analysis*, *63*(3), 194–199.

Glass, D. H. (2002). Coherence, explanation, and Bayesian networks. In M. O'Neill, R. F. E. Sutcliffe, C. Ryan, M. Eaton, & N. J. L. Griffith (Eds.), *Artificial Intelligence and Cognitive Science, 13th Irish Conference, AICS 2002* (p. 177–82). Berlin: Springer.

Glass, D. H. (2012). Inference to the best explanation: does it track truth? *Synthese*, *185*, 411–427.

Hahn, U., Harris, A. J., & Corner, A. (2016). Public reception of climate science: Coherence, reliability, and independence. *Topics in Cognitive Science*, *8*(1), 180–195.

Harris, A. J., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1366.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan.

Koscholke, J. (2016). Evaluating test cases for probabilistic measures of coherence. *Erkenntnis*, *81*(1), 155–181.

Koscholke, J., & Jekel, M. (2017). Probabilistic coherence measures: a psychological study of coherence assessment. *Synthese*, *194*(4), 1303–1322.

Koscholke, J., & Schippers, M. (2016). Against relative overlap measures of coherence. *Synthese*, *193*(9), 2805–2814.

Koscholke, J., Schippers, M., & Stegmann, A. (2019). New hope for relative overlap measures of coherence. *Mind*, *128*(512), 1261–1284.

Meijs, W. (2005). *Probabilistic Measures of Coherence*. Dissertation, Erasmus University, Rotterdam.

Meijs, W. (2006). Coherence as generalized logical equivalence. *Erkenntnis*, *64*(2), 231–252.

Olsson, E. J. (2002). What is the problem of coherence and truth? *Journal of Philosophy*, *99*(5), 246–72.

Olsson, E. J. (2021). Coherentist theories of epistemic justification. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 ed.). Metaphysics Research Lab, Stanford University.

Schippers, M. (2014). Probabilistic measures of coherence: From adequacy constraints towards pluralism. *Synthese*, *191*(16), 3821–3845.

Schupbach, J. N. (2011). New hope for Shogenji's coherence measure. *The British Journal for the Philosophy of Science*, *62*(1), 125–142.

Shogenji, T. (1999). Is coherence truth conducive? *Analysis*, *59*(4), 338–345.