New Hope for Relative Overlap Measures of Coherence

Jakob Koscholke
University of Hamburg
jakob.koscholke@uni-hamburg.de

Michael Schippers
University of Oldenburg
mi.schippers@uni-oldenburg.de

Alexander Stegmann
University of Oldenburg
alexander.stegmann@uni-oldenburg.de

Relative overlap measures of coherence have recently been shown to have two devastating properties: (i) according to the plain relative overlap measure, the degree of coherence of any set of propositions cannot be increased by adding further propositions, and (ii) according to the refined relative overlap measure, no set can be more coherent than its most coherent two-element subset. This result has been taken to rule out relative overlap as a foundation for a probabilistic explication of coherence. The present paper shows that this view is premature: we propose a relative overlap measure that does not fall victim to the two properties. The guiding idea is to employ a well-established recipe for the construction of coherence measures and to adapt it to the idea of relative overlap. We show that this new measure keeps up with, or even outperforms, former overlap measures in a set of desiderata for coherence measures and a collection of popular test cases. This result re-establishes relative overlap as a candidate for a proper formalization of coherence.

1. Introduction

It is a platitude among epistemologists that holding a coherent belief system is, *ceteris paribus*, better than holding an incoherent one. But regardless of whether one agrees with this view or not, the notion of coherence involved in this and similar platitudes has rarely been elucidated in a philosophically satisfying way. Some philosophers have employed phrases like 'hanging together', 'fitting together' or 'supporting each other' in order to characterize what they mean by calling a set of propositions coherent (see Olsson 2005). But in the early days of coherentism, even prominent advocates of this position have already complained about its fragile conceptual foundation. For

Mind, Vol. 00 . 0 . 2018
doi:10.1093/mind/fzy037

instance, Rescher has noted, 'the coherence theorists themselves have not always been too successful in explicating the nature of coherence' (Rescher 1973, p. 33). Even more pessimistically, after having spent years of research on coherentist theories of justification, BonJour has remarked, 'the precise nature of coherence remains a largely unsolved problem' (BonJour 1999, p. 124).

Of course, not only proponents of coherentism have discussed the lack of precision regarding the concept of coherence. As Douven and Meijs have pointed out, many opponents have objected that 'the notion of coherence is hopelessly vague' (Douven and Meijs 2007, pp. 405–6). Or as Bovens and Hartmann have put it, 'There is a long-standing embarrassment here. A definition of what it means for one set of propositions to be more coherent than another set has not been forthcoming' (Bovens and Hartmann 2003b, p. 602). Over the years, however, there have been a number of attempts to formalize the notion of coherence using the tools provided by probability theory. The results are so-called probabilistic measures of coherence; for an overview see Schippers (2014b). Such measures are functions assigning real numbers to sets of propositions relative to some joint probability function—ideally, according to how strongly the propositions cohere. These measures are typically subdivided into the following three classes: measures based on *deviation from independence* (Shogenji 1999; Schupbach 2011; Koscholke 2016), on *average mutual support* (Fitelson 2003; Douven and Meijs 2007; Roche 2013; Schippers 2014b), and on *relative overlap* (Glass 2002; Olsson 2002; Meijs 2006).

In this paper, we concentrate on the third class of proposals. This class has recently received some attention in the literature, since it has been shown that existing relative overlap measures have two properties which can be considered to disqualify them as proper measures of coherence. More specifically, it has been shown in Koscholke and Schippers (2015) that (i) according to the simple relative overlap measure proposed by Glass (2002) and Olsson (2002), the degree of coherence of any set of propositions cannot be increased by adding propositions, and that (ii) according to the subset-sensitive relative overlap measure proposed by Meijs (2006), no set can be more coherent than its most coherent two-element subset. One of the conclusions that has been drawn from this result is that the idea of relative overlap is not suitable for providing a formal explication of the concept of coherence. In this paper, however, we present a relative overlap measure that does not fall victim to the two aforementioned properties. Even more, we show that this new measure satisfies a set of coherence-related desiderata and performs very well in a collection of test cases from the literature. If our reasoning is correct, this result re-establishes relative overlap as a candidate for a proper formalization of coherence.

The paper is structured as follows: in §2 we introduce the two existing relative overlap measures and the results presented against them; in §3 we present the new relative overlap measure together with the main result showing that this measure does not fall victim to the two problematic properties discussed; in §4 we investigate some further properties of this new measure, and in §5 we examine its performance in a well-established set of test cases; finally, in §6 we summarize our results. To have a uniform framework in the following, let a probabilistic coherence measure be any function $C : \mathbf{L} \times \mathbf{P} \to \mathbb{R}$, where $\mathbf{L}$ is the set of all non-empty, non-singleton subsets of some propositional language L, and $\mathbf{P}$ is the set of probability functions P that can be defined over L. Accordingly, the real numbers assigned to some pair $(X, P) \in \mathbf{L} \times \mathbf{P}$ by C are supposed to represent degrees of coherence of the set X consisting of the propositions $x_1,\ldots,x_n$ under the joint probability distribution P. Given this general characterization of a coherence measure, let us now turn to the measures and the arguments put forward against them.

2. Against relative overlap

According to a fairly popular account, the degree to which a finite number of propositions cohere can be understood as the size of their overlapping set-theoretic surface compared to the size of their total set-theoretic surface. In probabilistic terms, this corresponds to the probability that all of the set's propositions are true together compared to the probability that at least one of them is. Accordingly, if the propositions are equivalent, their overlapping surface is identical to their total surface, and hence they are judged maximally coherent according to this account. By contrast, if they are pairwise inconsistent, there is no overlap at all, and hence they are considered minimally coherent according to this account. A corresponding measure has been proposed independently by Glass (2002) and Olsson (2002):

$$\mathcal{O}(X, P) = \frac{P\left(\bigwedge_{x_i \in X} x_i\right)}{P\left(\bigvee_{x_i \in X} x_i\right)}$$

Although this idea of formalizing coherence seems quite natural, it comes with a very counter-intuitive property: it violates a rather weak but appealing desideratum concerning the notion of coherence. This desideratum can already be found in a remark by Meijs:

> [A] general intuition about coherence is the idea that it must be possible to increase the coherence of an information set by adding propositions. Evidently, it will not always be the case that adding a proposition increases a set's coherence from an intuitive standpoint. For example, if we add a proposition that is inconsistent with the conjunction of the propositions in a set, then we should not expect coherence to increase. (Meijs 2007, p. 161)

Let us sidestep here the question of which features the added propositions must have in order to increase or decrease coherence. What is important for now is that it should be *possible* to increase or decrease coherence. And also, Meijs and many others will certainly agree that by adding propositions it should also be possible that coherence remains unchanged. Summing up: it should be possible to increase, decrease or leave unchanged the degree of coherence of some set of propositions by adding a number of arbitrary propositions. Formally, this amounts to the following condition:

Extension. For each of the following three cases there is a pair $(X, P) \in \mathbf{L} \times \mathbf{P}$ such that if $X \subset X'$, then:

(1) $C(X', P) > C(X, P)$

(2) $C(X', P) = C(X, P)$

(3) $C(X', P) < C(X, P)$

This condition is satisfied by a coherence measure if for each of the three cases there exists some set under some probability function such that the measure exhibits the corresponding behaviour. As can easily be shown, the condition is trivially satisfied by the majority of coherence measures discussed in the literature. It can therefore be considered quite surprising that the following holds (Koscholke and Schippers, 2015):

Theorem 1.
$\mathcal{O}$ violates case 1 of *Extension* in general, and additionally violates case 2 of *Extension* for any non-zero probability function.

This result provides a powerful argument against $\mathcal{O}$: a measure behaving this way simply trivializes the whole idea of *measuring* coherence. A coherence measure according to which extending a set of propositions cannot increase its degree of coherence is of little or no epistemological value. A detailed proof of this result can be found in Koscholke and Schippers (2015). However, there is a shorter and quite interesting version to which we would like to draw the reader's attention: the short version of the proof is that $\mathcal{O}$ is actually and rather surprisingly a probability function, namely, the following: $P(x_1 \wedge \ldots \wedge x_n | x_1 \vee \ldots \vee x_n)$. Since probability functions satisfy the well-known Boole-Fréchet inequalities (Fréchet, 1935), the values provided by $\mathcal{O}$ cannot be higher if the number of propositions increases, and can only remain identical under certain conditions. Therefore, our lesson from this result becomes the following: $\mathcal{O}$ should not be considered a proper measure of coherence, because it is a probability function. It is, of course, also possible for a coherence measure to violate the aforementioned property for other reasons. But we consider it a valuable insight that being a probability function is sufficient for not being a proper coherence measure—if one considers the aforementioned condition plausible, of course. This does not, however, mean that a measure *based on* a probability function cannot be a proper coherence measure. In fact, we are going to propose such a measure in §3 of this paper.

A quite prominent example where a violation of case 1 of the aforementioned condition occurs is Bovens and Hartmann's (2003a) well-known Tweety case: suppose that by independent and equally reliable sources one is provided with two pieces of information about someone's pet Tweety, namely, that Tweety is a bird (x_1) and that Tweety is a ground-dweller (x_2). Later, one also receives the piece of information that Tweety is a penguin (x_3). Bovens and Hartmann have provided the following joint probability distribution to model the situation: $P(x_1 \wedge x_2 \wedge x_3) = P(\neg x_1 \wedge \neg x_2 \wedge \neg x_3) = 1/100$, $P(x_1 \wedge \neg x_2 \wedge \neg x_3) = P(\neg x_1 \wedge x_2 \wedge \neg x_3) = 49/100$. According to Bovens and Hartmann, the set $\{x_1, x_2\}$ should not be considered very coherent, because birds are usually conceived of as flying—intuitively, there seems to be some tension between the propositions x_1 and x_2. However, adding the third proposition x_3 relieves this tension. Hence the extended set $\{x_1, x_2, x_3\}$ should be considered more coherent than the original set. Still, $\mathcal{O}$ fails to capture this intuitive verdict: $\mathcal{O}(\{x_1, x_2\}, P) = \mathcal{O}(\{x_1, x_2, x_3\}, P) \approx 0.01$. Given Theorem 1, this result is not very surprising: there are *no* situations in which extending a set yields a higher $\mathcal{O}$-value.

One might try to relativize this result by pointing out that there is a refined version of Glass's and Olsson's relative overlap measure proposed by Meijs (2006). And in fact this refined measure provides the right verdict in the Tweety case, as Meijs himself has shown. The basic idea underlying this measure is that coherence is subset-sensitive: in order to determine the degree of coherence of some set relative to some probability function, one should take into account the degrees of coherence of all non-empty, non-singleton subsets X' of the target set X and take the straight average over these values. For a set consisting of n propositions, the corresponding number of such subsets is $m = (2^n - n) - 1$. The measure thus reads as follows:

$$\mathcal{O}'(X, P) = m^{-1} \sum_{i=1}^{m} \mathcal{O}(X', P)_i$$

Now, the first thing one might be interested in is whether this modified measure, apart from satisfying a special case of the aforementioned condition, namely, the Tweety case, also satisfies the aforementioned condition in general. In fact it does:

Theorem 2.
$\mathcal{O}'$ satisfies *Extension*.

The proof of this statement is straightforward: for each of the three cases of the condition one simply has to provide a suitable example consisting of a set of propositions under some probability function. Somewhat ironically, the Tweety case which has previously been employed as an example against the regular overlap measure $\mathcal{O}$, works as a positive example for the refined measure $\mathcal{O}'$. As regards case 1 of the condition, it can be seen that $\mathcal{O}'(\{x_1, x_2\}, P) \approx 0.01 < \mathcal{O}'(\{x_1, x_2, x_3\}, P) \approx 0.015$. As an example for case 2 of the condition, consider a very simple joint probability distribution over three arbitrary propositions y_1, y_2 and y_3 such that $P(y_1 \wedge y_2 \wedge y_3) = 1$. Quite obviously, this yields $\mathcal{O}'(\{y_1, y_2\}, P) = \mathcal{O}'(\{y_1, y_2, y_3\}, P) = 1$. Finally, as a proof for case 3 of the condition, simply consider a joint probability distribution over the propositions z_1, z_2 and z_3 such that all eight Boolean combinations receive the same probability of 1/8. It can easily be seen that $\mathcal{O}'(\{z_1, z_2\}, P) \approx 0.33 > \mathcal{O}'(\{z_1, z_2, z_3\}, P) \approx 0.28$.

The refined overlap measure $\mathcal{O}'$ might be able to deal with the Tweety case, and even satisfy the set extension condition. But there is another condition which causes difficulties for this measure. These difficulties are due to the simple fact that $\mathcal{O}'$ is just a weighted average over $\mathcal{O}$-values and therefore inherits some variations of the properties of $\mathcal{O}$. The critical condition runs as follows:

Unboundedness. There is a $(X, P) \in \mathbf{L} \times \mathbf{P}$ such that $C(X, P) > C(X', P)$ for any $X' \subset X$ with $|X'| = 2$.

This condition is different from the condition discussed previously, but the underlying idea is closely related. The basic idea is that the most coherent two-element subset should not limit the degree of coherence of the target set. A violation of this condition trivializes the idea of measuring coherence in a different way: if it is always the case for a coherence measure that the most coherent two-element subset limits the degree of coherence of the target set, the degrees of coherence of the remaining sets do not really seem to matter—but intuitively they do. Again, just as for the condition discussed previously, it is worth noticing that most coherence measures proposed in the literature satisfy this condition. But just like the simple measure $\mathcal{O}$, the refined measure $\mathcal{O}'$ does not:

Theorem 3.
$\mathcal{O}'$ violates *Unboundedness*.

For a rigorous proof of this statement, the reader is again referred to Koscholke and Schippers (2015). Interestingly, the proof not only shows that the result holds for the measure $\mathcal{O}'$, which is based on the straight average over the relevant $\mathcal{O}$-values; the result holds for any averaging procedure. Hence it is impossible for $\mathcal{O}'$ and its variations to judge a set more coherent than it judges the set's most coherent two-element subset. One class of cases where this property might be considered problematic is that class of cases where the propositions to be evaluated are tied together by inferential relations such as deductive entailment, evidential support, or explanation. For instance, consider the set $\{x_1, x_1 \rightarrow x_2, x_2\}$, and keep in mind that the arguments of probabilistic coherence measures are usually not assumed to be closed under *modus ponens*. It seems that any of the two-element subsets $\{x_1, x_1 \rightarrow x_2\}$, $\{x_1, x_2\}$ or $\{x_1 \rightarrow x_2, x_2\}$ can be made more coherent by adding the corresponding missing proposition, since this establishes strong inferential relations within the set that are usually considered coherence-increasing. But even if one does not find this argument convincing, maybe due to the connectives involved, there is another argument which does not rely on this aspect. Consider some arbitrary set of one million propositions—should it be impossible for a coherence measure to judge such a set more coherent than its most coherent two-element subset? The refined overlap measure $\mathcal{O}'$ has to answer this question positively. But intuitively, there is a plethora of subsets with more than two elements that could be more coherent. Hence there seems to be no good reason for this general type of behaviour.

To summarize the results discussed in this section: neither the simple nor the refined overlap measure seems to be suitable to capture the intuitive notion of coherence. This is because they violate two very simple and appealing coherence-related conditions. One might therefore be inclined to think that the idea of coherence as relative overlap is misguided in general. But this conclusion would be premature. In fact, the next section is devoted to showing that it is false.

3. New hope for relative overlap

The aforementioned results have been interpreted as suggesting that the idea of relative overlap is simply inadequate as a foundation for a proper probabilistic measure of coherence. But this conclusion seems rather strong, especially since there is no rigorous characterization of what a relative overlap measure actually is—apart from the measures $\mathcal{O}$ and $\mathcal{O}'$, which are species of this class, of course. It should be obvious, however, that as long as it is not clear what a relative overlap measure is, the claim that such measures cannot be probabilistic measures of coherence is built on sand. In fact, the claim can easily be refuted: it is enough to present a measure which counts as a relative overlap measure but does not have the problems affecting $\mathcal{O}$ and $\mathcal{O}'$. This is the strategy we are going to pursue in this section.

As mentioned before, measures based on relative overlap form one of the three classes of extant probabilistic measures of coherence. Another quite prominent class of measures is based on the idea of coherence as average mutual support. This class is based on a general recipe developed by Douven and Meijs (2007). The basic idea runs as follows: in order to compute the degree of coherence of a set X, first consider all pairs (X', X''), where X' and X'' are non-empty, disjoint subsets of X; notice that for n propositions there are exactly $(3^n - 2^{n+1}) + 1$ such pairs (see Roche 2013); then, for each pair, take the conjunction over the propositions in the respective set, compute the degree of support according to some chosen probabilistic measure of support, S and calculate the arithmetic mean over the resulting values. Since there is a plethora of probabilistic measures of support to choose from (for an overview, see Crupi, Tentori and Gonzalez 2007), a huge variety of potential coherence measures can be generated with this recipe.

For our new measure, we are going to make quite an unorthodox choice for S: we will simply use $\mathcal{O}$. So far, only incremental or absolute measures of confirmation have been used, and $\mathcal{O}$ is clearly neither. But already at this point, choosing this measure has a benefit: unlike many probabilistic measures of support, the measure $\mathcal{O}$ is commutative for pairs of propositions. This helps us to reduce computational costs, since the number of $\mathcal{O}$-values that have to be taken into account in the calculations collapses to $l = [(3^n - 2^{n+1}) + 1]/2$. However, even if we do take all redundant values into account, this is not too bad: the results will be the same, owing to the averaging procedure. The resulting measure thus reads:

$$\mathcal{O}^*(X, P) = l^{-1} \sum_{i=1}^{l} \mathcal{O}\left(\{ \bigwedge_{x_j \in X'} x_j, \bigwedge_{x_k \in X''} x_k \}, P \right)_i$$

The idea of using the regular relative overlap measure as an ingredient in Douven and Meijs's recipe can be referred to termed coherence as *average mutual relative overlap*—or for gourmets: *relative overlap à la Douven and Meijs*. Now that we are equipped with this new type of relative overlap measure, we are first interested in how it performs with respect to the conditions in which the other relative overlap measures $\mathcal{O}$ and $\mathcal{O}'$ have failed. Fortunately, this is not too difficult to examine, and we can therefore state the following main result:

Theorem 4.
$\mathcal{O}^*$ satisfies *Extension* and *Unboundedness*.

Let us first focus on the claim that the new measure satisfies the second of the two conditions, namely, the unboundedness condition. Interestingly, the probability distribution provided by the Tweety case is sufficient for the proof of this claim. Based on the pairwise $\mathcal{O}$-values presented in the left-hand graph in Fig. 1, it is quite easy to calculate that $\mathcal{O}^*(\{x_1, x_2, x_3\}, P) \approx (4 \times 0.02 + 0.01 + 1)/6 \approx 0.18$. Moreover, it does not take much effort to see that this value is larger than the $\mathcal{O}^*$-values of any of the two-element subsets: $\mathcal{O}^*(\{x_1, x_2\}, P) \approx 0.01$, $\mathcal{O}^*(\{x_1, x_3\}, P) \approx 0.02$, and $\mathcal{O}^*(\{x_2, x_3\}, P) \approx 0.02$. This result is especially interesting, because unlike the other measures $\mathcal{O}$ and $\mathcal{O}'$, the new

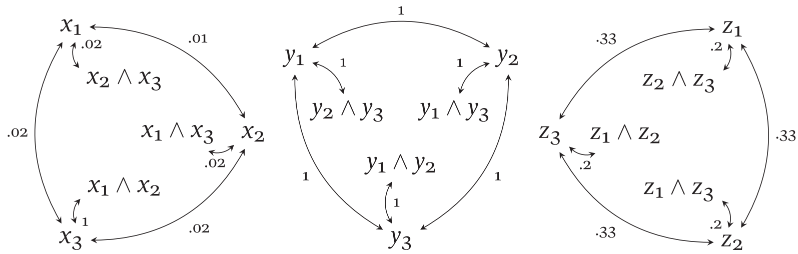

Figure 1: Pairwise $\mathcal{O}$-values for Theorem 4

measure $\mathcal{O}^*$ judges the extended set *much more coherent* than the original set. This nicely captures the intuition that the information that Tweety is a penguin relaxes, to quite a significant degree, the tension between the information that Tweety is a bird and the information that Tweety does not fly.

Let us now turn to the claim that the new measure also satisfies the first of the two conditions, namely, the set extension condition. Here the Tweety case comes in handy again: it is sufficient to show that the measure satisfies case 1 of the condition, since, as we have already noted, $\mathcal{O}^*(\{x_1, x_2\}, P) \approx 0.01 < \mathcal{O}^*(\{x_1, x_2, x_3\}, P) \approx 0.18$. As a proof for case 2 of the condition, consider again a very simple joint probability distribution over y_1, y_2 and y_3, such that $P(y_1 \wedge y_2 \wedge y_3) = 1$. As can be seen in the middle graph in Fig. 1, this yields $\mathcal{O}^*(\{y_1, y_2\}, P) = \mathcal{O}^*(\{y_1, y_2, y_3\}, P) = 1$. Finally, for case 3 of the condition, consider again the probability distribution over three propositions, z_1, z_2 and z_3, such that all eight Boolean combinations receive the same probability of 1/8. Some simple calculations based on the $\mathcal{O}$-values presented in the right-hand graph in Fig. 1 yield $\mathcal{O}^*(\{z_1, z_2\}, P) \approx 0.33 > \mathcal{O}^*(\{z_1, z_2, z_3\}, P) \approx 0.265$.

It seems that our strategy has worked out: the new measure is true to the idea of coherence as relative overlap but satisfies the two conditions the two previous relative overlap measures have violated—and quite ironically, this can be shown using a test case that has been employed by other authors to argue against relative overlap measures of coherence. Nevertheless, our investigation should not end here. In the next section, we shall examine some further properties of the new measure and the relationships between the measures discussed.

4. Further properties of $\mathcal{O}^*$

It is interesting to see that the new measure does not fall victim to the problems the other two relative overlap measures have. But what is even more interesting is that it also satisfies a number of desiderata that have been proposed for probabilistic measures of coherence (for an overview, see Schippers 2014b). The main motivation underlying these desiderata is to describe how the concept of coherence is related to other philosophically interesting notions.

One such desideratum has been proposed by Fitelson (2003), as well as Siebel and Wolff (2008). The basic idea is that cases of logically equivalent propositions should be considered paradigmatic examples of maximal coherence. More precisely:

Equivalence. For any $(X, P) \in \mathbf{L} \times \mathbf{P}$ such that all $x_i, x_j \in X$ are logically equivalent: $C(X, P) = \max(C)$.

It is not difficult to see that $\mathcal{O}^*$ satisfies this condition. By definition, logically equivalent propositions can only be true or false together. Set-theoretically speaking, this means that the propositions' overlapping surface is identical to their total surface. And since relative overlap $\mathcal{O}$ is nothing but the ratio between these the two quantities, its value must be 1 and thus maximal—provided that the propositions are satisfiable, of course. Finally, since $\mathcal{O}^*$ is entirely defined by pairwise $\mathcal{O}$-values, the resulting value must also be 1 and thus maximal.

A closely related desideratum has also been proposed by Fitelson (2003) and investigated more thoroughly by Schippers (2014a). This desideratum can be considered the natural counterpart to the aforementioned desideratum: if logical equivalence is taken to be the paradigmatic case of maximal coherence, logical inconsistency should be considered the paradigmatic case of minimal coherence—or maximal incoherence, if you will. The corresponding condition reads as follows:

Inconsistency. For any $(X, P) \in \mathbf{L} \times \mathbf{P}$ such that all $x_i, x_j \in X$ are individually satisfiable but jointly logically inconsistent: $C(X, P) = \min(C)$.

The proof that $\mathcal{O}^*$ satisfies this condition is analogous to the sketch presented earlier. By definition, logically inconsistent propositions cannot be true together. Set-theoretically speaking, this means that the propositions' overlapping surface is empty and thus relative overlap must be 0, which is the minimal value. Since $\mathcal{O}^*$ is defined in terms of pairwise $\mathcal{O}$-values, the resulting value must also be 0 and thus minimal.

The next desideratum has been proposed by Bovens and Olsson (2000) and generalized by Schippers (2014b). The underlying intuition is that increasing conditional probabilities among the propositions in some set should increase its coherence. More precisely, suppose there is some set X over which two joint probability functions are defined. If under the first all relevant conditional probabilities are higher than under the second, then the degree of coherence under the first should also be higher than under the second. Formally:

Agreement. Let $P_1, P_2 \in \mathbf{P}$ be such that for all non-empty, disjoint subsets X', X'' of $X \in \mathbf{L}$, the following inequality holds:

$$P_1\left(\bigwedge_{x_i \in X'} x_i | \bigwedge_{x_j \in X''} x_j\right) > P_2\left(\bigwedge_{x_i \in X'} x_i | \bigwedge_{x_j \in X''} x_j\right)$$

Then it also holds that $C(X, P_1) > C(X, P_2)$.

It is slightly more difficult to prove that $\mathcal{O}^*$ satisfies this condition. However, we can employ a result shown by Glass (2005). For two propositions, the following relationship holds:

$$\mathcal{O}(\{x_1, x_2\}, P) = \frac{P(x_1|x_2) \times P(x_2)}{P(x_1) + P(x_2) - P(x_1|x_2) \times P(x_2)}$$
$$= \left(\frac{1}{P(x_1|x_2)} + \frac{1}{P(x_2|x_1)} - 1\right)^{-1}$$

Since $\mathcal{O}^*$ is defined in terms of pairwise $\mathcal{O}$-values, the proof almost comes for free: if we compare a set of propositions under two probability functions such that under the first all relevant conditional probabilities are higher than under the second, then by the equation given above, the corresponding pairwise $\mathcal{O}$-values will also be higher.

Moreover, through this result we obtain another result for free. Fitelson (2003) has proposed a desideratum which draws a connection between coherence and probabilistic dependence. The basic idea is that a set of propositions where each subset is independent should be assigned some threshold value indicating neither coherence nor incoherence. Correspondingly, cases of positive dependence or negative dependence should be rewarded with values higher or lower than this threshold value. More precisely:

Dependence. There is a threshold θ for C such that for any $(X, P) \in \mathbf{L} \times \mathbf{P}$:

(1) $C(X, P) > \theta$ if all $X' \subseteq X$ are positively dependent under P.

(2) $C(X, P) = \theta$ if all $X' \subseteq X$ are independent under P.

(3) $C(X, P) < \theta$ if all $X' \subseteq X$ are negatively dependent under P.

Let us turn to the free result now. First, some background information. As some readers will know, Schippers (2014b) has proved a surprising impossibility result to the effect that, for purely mathematical reasons, there can be no probabilistic coherence measure which satisfies both Bovens and Olsson's desideratum on coherence and conditional probabilities and Fitelson's desideratum on coherence and probabilistic dependence. More precisely:

Theorem 5
Agreement and *Dependence* are inconsistent.

Hence, thanks to Theorem 5 and the fact that $\mathcal{O}^*$ satisfies the former desideratum, we obtain the free result that $\mathcal{O}^*$ does not satisfy the latter desideratum. Now, one might be inclined to think that this is bad news for the new measure. After all, Fitelson's idea that coherence is connected to probabilistic dependence seems quite appealing. But this would be premature. Bovens and Olsson's idea that higher mutual conditional probabilities entail higher degrees of coherence seems as appealing as the aforementioned—there is not a single argument in the literature indicating that one could be preferred over the other. Accordingly, the conclusion that has been drawn from Theorem 5 is that we should embrace pluralism concerning probabilistic measures of coherence. In other words, we must give up the idea that there is *one* true measure of coherence that satisfies all desiderata. Instead, Theorem 5 suggests that there are at least two distinct classes of coherence measures representing two different facets of the concept of coherence. Therefore, future research should focus on finding the most promising members of each of these two classes. The implication for the new measure: future research should compare $\mathcal{O}^*$ only with measures which also satisfy Bovens and Olsson's desideratum, rather than with measures belonging to the disjoint class of measures satisfying Fitelson's desideratum. To do otherwise would be to compare apples with oranges. Let us summarize the results of this section by stating the following:

Theorem 6
$\mathcal{O}^*$ satisfies *Equivalence*, *Inconsistency* and *Agreement*, but not *Dependence*.

This result closes our investigation of further properties of the new relative overlap measure. It is quite interesting to see that the new measure circumvents the problems the other two measures suffered from, but still satisfies some well-established coherence desiderata. In the next section, we will examine the measure's performance in a number of test cases discussed in the literature.

5. Test cases

We have already examined how $\mathcal{O}^*$ performs in the Tweety case. But there are more test cases for probabilistic coherence measures in the literature (for an overview, see Koscholke 2015). Since both $\mathcal{O}$ and $\mathcal{O}'$ are known to master quite a number of them, it would be interesting to see how $\mathcal{O}^*$ performs. For readers who are afraid of the computational effort, we have some good news: we do not have to calculate all values for the new measure again. It is easy to see that for cases of two propositions, the values of $\mathcal{O}$, $\mathcal{O}'$ and $\mathcal{O}^*$ are identical. This concerns the majority of test cases proposed in the literature, in particular, Akiba's (2000) die case, Bovens and Hartmann's (2003a) Tokyo murder case, Siebel's (2004) pickpocketing robber case, Glass's (2005) dodecahedron case, Meijs and Douven's (2005) plane lottery case, Meijs's (2006) albino rabbit case and Douven and Meijs's (2007) Samurai sword case.

Still, the fact that we do not have to calculate the values for these cases does not mean that we should not appreciate the results. In fact, as one can verify by consulting Koscholke (2015), the measures $\mathcal{O}$ and $\mathcal{O}'$, and hence the new measure $\mathcal{O}^*$, perform extremely well: they master every single test case except Akiba's (2000) die case. However, as Shogenji (2001) and Olsson (2005, p. 101) have convincingly argued, there are good reasons to believe that the intuitive coherence assessment provided by Akiba's test case is wrong: Akiba has claimed that the proposition x_1, that a fair die will land on 2, is as coherent with the proposition x_2, that it will land on 2 or 4, as it is with the proposition x_3, that it will land on 2, 4 or 6. The reason for his view is that both x_2 and x_3 are logical consequences of x_1. But he seems to ignore that x_2 is logically stronger than x_3, and hence the agreement regarding the number on which the die will land is stronger in the set $\{x_1, x_2\}$ than in the set $\{x_1, x_3\}$. If we take the correct intuition to be that the first set should be more coherent than the second, then the new measure $\mathcal{O}^*$ masters Akiba's case. This leaves us with a manageable set of four test cases, including BonJour's (1985) raven case, Bovens and Hartmann's (2003a) culprit case, Schupbach's (2011) robber case and Fitelson's (2015) inconsistent testimony case.

Let us begin with BonJour's (1985) famous raven case. In his seminal *The Structure of Empirical Knowledge* BonJour presented this example to demonstrate the difference between a set of coherent and a set of less coherent propositions. The first, coherent, set consists of three propositions: x_1, that all ravens are black; x_2, that some randomly chosen bird is a raven; and x_3, that this randomly chosen bird is black. The second, less coherent, set also consists of three propositions: y_1, that some randomly chosen chair is brown; y_2, that electrons are negatively charged; and y_3, that today is Thursday. Intuitively, the first set is much more coherent than the second set, since the propositions in the second set seem completely unrelated. Bovens and Hartmann (2003a) have provided two suitable probability distributions to model these two sets. They are shown in Fig. 2. Using the information provided, we can verify that all considered measures judge the first set more coherent than the second, that is, $\mathcal{O}(\{x_1, x_2, x_3\}, P) \approx 0.108 > \mathcal{O}(\{y_1, y_2, y_3\}, P) \approx 0.027$, $\mathcal{O}'(\{x_1, x_2, x_3\}, P) \approx 0.176 > \mathcal{O}'(\{y_1, y_2, y_3\}, P) \approx 0.114$ and $\mathcal{O}^*(\{x_1, x_2, x_3\}, P) \approx 0.212 > \mathcal{O}^*(\{y_1, y_2, y_3\}, P) \approx 0.098$. One interesting aspect to which we would like to draw the reader's attention is that the difference between the values $\mathcal{O}^*$ assigns to both sets is much larger than for the two other measures: 0.081 for $\mathcal{O}$, 0.062 for $\mathcal{O}'$, but 0.114 for $\mathcal{O}^*$. This captures the intuition that it is not only the case that the first set is more coherent than the second—it is *much* more coherent. Moreover, it is quite interesting to see that although this case is based on the intuition that probabilistic dependence is somehow relevant for coherence—the propositions are positively dependent in the first case but independent in the second—all overlap measures master BonJour's raven case.

Let us proceed to the next case. Bovens and Hartmann (2003a) have presented a huge variety of test cases for probabilistic measures of coherence in their *Bayesian Epistemology*. One of them is their culprit case. It runs as follows. Suppose we have to identify a culprit in a murder case. Consider the first situation in which one is provided with

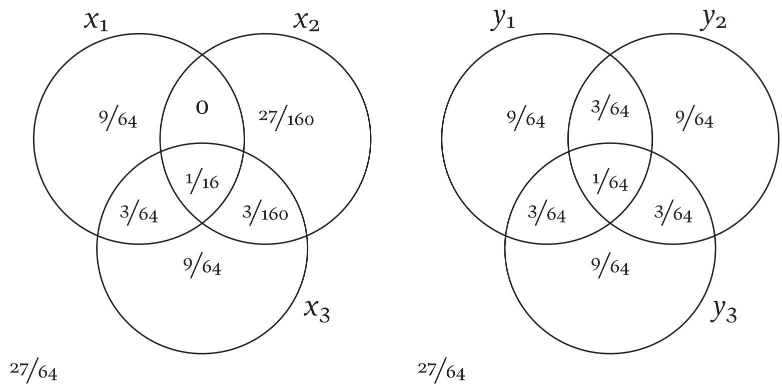

Figure 2: Distributions for the raven case

the following reports from independent and equally reliable sources: x_1, that the culprit was a woman; x_2, that the culprit had a Danish accent; and x_3, that the culprit drove a Ford. Analogously, in the second situation one is provided with the following reports: x_1, that the culprit wore Coco Chanel shoes; x_2, that the culprit had a French accent; and x_3, that the culprit drove a Renault. The corresponding probability distributions are shown in Fig. 3. It can already be seen that in the second situation there is perfect relative overlap between the three reports. By contrast, in the first situation the overlapping surface is identical to the second situation, but each pairwise overlap is larger than the joint overlap and each non-overlapping surface is even larger than each pairwise and joint overlap. Accordingly, Bovens and Hartmann have argued that the second set should be judged more coherent than the first. The measures behave accordingly: they all assign a maximal value of 1 to the second set. On the other hand, we have $\mathcal{O}(\{x_1, x_2, x_3\}, P) \approx 0.082$, $\mathcal{O}'(\{x_1, x_2, x_3\}, P) \approx 0.208$ and $\mathcal{O}^*(\{x_1, x_2, x_3\}, P) \approx 0.190$. It is quite interesting to see that although Bovens and Hartmann have proposed their very own approach to modelling coherence probabilistically, their test case relies heavily on the intuition that relative overlap is relevant for coherence.

Another test case has been proposed by Schupbach (2011). It has been employed to point out a difficulty for Shogenji's (1999) well-known measure of coherence based on the idea of deviation from independence. The case runs as follows: there are eight equiprobable suspects for a robbery, out of whom one is the robber for sure. Independent and equally reliable witness reports are given. Now, consider the following two situations. In the first situation, the witnesses

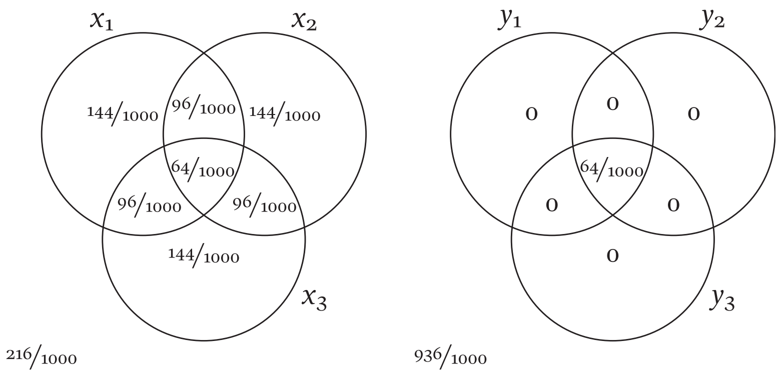

Figure 3: Distributions for the culprit case

reports are: x_1, that it was suspect 1, 2 or 3; x_2, that it was suspect 1, 2 or 4; and x_3, that it was suspect 1, 3 or 4. In the second situation, the witnesses reports are: y_1, that it was suspect 1, 2 or 3; y_2, that it was suspect 1, 4 or 5; and y_3, that it was suspect 1, 6 or 7. This fully determines two joint probability distributions corresponding to the two situations. They are shown in Fig. 4. As Schupbach has convincingly argued, the reports are more coherent in the first than in the second situation, because in the first situation each pair of reports agrees upon two suspects, whereas in the second situation it is only one suspect they agree upon. Given these probabilities, the measures assign the following values: $\mathcal{O}(\{x_1, x_2, x_3\}, P) \approx 0.25 > \mathcal{O}(\{y_1, y_2, y_3\}, P) \approx 0.143$, $\mathcal{O}'(\{x_1, x_2, x_3\}, P) \approx 0.438 > \mathcal{O}'(\{y_1, y_2, y_3\}, P) \approx 0.186$ and $\mathcal{O}^*(\{x_1, x_2, x_3\}, P) \approx 0.375 > \mathcal{O}^*(\{y_1, y_2, y_3\}, P) \approx 0.267$. It is again interesting to see that just like the aforementioned test case, this case is also based on a relative overlap intuition. Based on this observation, one might argue that the selection of discussed test cases is somewhat biased. To a certain extent, we agree with this. However, the fact that there are several cases based on this intuition might also indicate that relative overlap simply plays a key role in assessments of coherence.

Our final test case is due to Schippers and Siebel (2015), who have investigated how probabilistic coherence measures fare in situations involving inconsistent sets of propositions. One might ask whether or not it makes any sense at all to assign degrees of coherence to inconsistent sets, but Schippers and Siebel have something more subtle in mind: even if a set is logically inconsistent, it can be so in quite different ways—it can be inconsistent because all the propositions are jointly

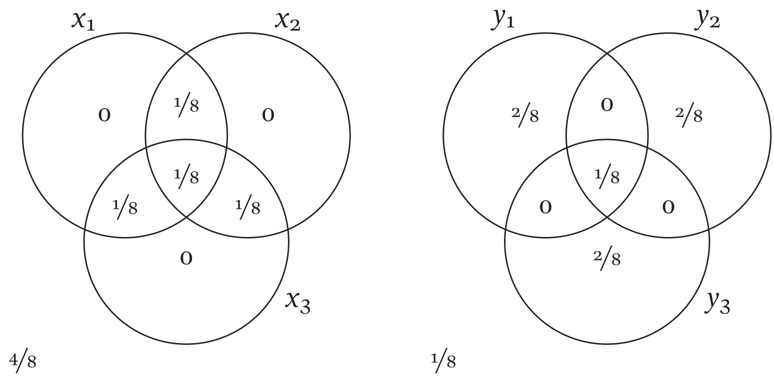

Figure 4: Distributions for the robber case

inconsistent or because they are pairwise inconsistent. The latter entails the former but not vice versa. Hence a set consisting of pairwise inconsistent propositions could be considered less coherent than a set of pairwise consistent but jointly inconsistent propositions. Accordingly, their test case runs as follows. There are eight equiprobable suspects for a robbery and exactly one of them is the robber. In the first situation, the reports are: x_1, that it was suspect 1 or 2; x_2, that it was suspect 2 or 3; and x_3, that it was suspect 1 or 3. In the second situation, the reports are: y_1, that it was suspect 1 or 2; y_2, that it was suspect 3 or 4; and y_3, that it was suspect 5 or 6. The corresponding probability distributions are shown in Fig. 5. The results in this test case are particularly interesting, since they allow us to rule out one measure, namely, $\mathcal{O}$: it assigns 0 to both sets. By contrast, $\mathcal{O}'(\{x_1, x_2, x_3\}, P) \approx 0.25 > \mathcal{O}'(\{y_1, y_2, y_3\}, P) = 0$ and $\mathcal{O}^*(\{x_1, x_2, x_3\}, P) \approx 0.167 > \mathcal{O}^*(\{y_1, y_2, y_3\}, P) = 0$. It is interesting to see that $\mathcal{O}^*$ is slightly more conservative when it comes to the degree of coherence for the first set. This matches our intuition that although the first set is more coherent than the second, it is still not very coherent—simply because, just like the second set, it contains propositions that cannot be true together.

One detail that might have caught the reader's attention is that the preceding test case results are very similar: the three measures rank-order almost any pair consisting of sets of propositions identically. This might raise the question of ordinal equivalence between the measures, that is, the question whether there are pairs of sets of propositions for which two measures can give different rank-orderings. Since we have already noticed that in the case of two propositions the measures are identical and hence ordinally equivalent, we must

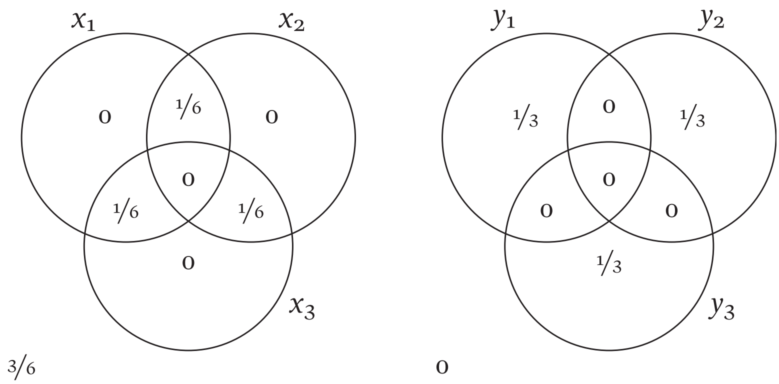

Figure 5: Distributions for the inconsistent testimony case

look for counterexamples in the space of sets containing at least three propositions. Fortunately, the Tweety case is just such a case: it shows that the pairs $(\mathcal{O}, \mathcal{O}')$ and $(\mathcal{O}, \mathcal{O}^*)$ are not ordinally equivalent, because $\mathcal{O}$ rank-orders the sets $\{x_1, x_2\}$ and $\{x_1, x_2, x_3\}$ identically, whereas the other two measures rank-order the first below the second. And the Tweety case turns out to be even more versatile: it also provides a proof that the pair $(\mathcal{O}', \mathcal{O}^*)$ is not ordinally equivalent. So far, we have focused on adding the proposition x_3, that Tweety is a penguin, to the set consisting of the proposition x_1, that Tweety is a bird, and x_2, that Tweety is a ground-dweller. But what about starting with the set $\{x_1, x_3\}$ and extending it to $\{x_1, x_2, x_3\}$? Our calculations yield $\mathcal{O}'(\{x_1, x_3\}, P) = 0.02 > \mathcal{O}'(\{x_1, x_2, x_3\}, P) \approx 0.015$, whereas $\mathcal{O}^*(\{x_1, x_3\}, P) = 0.02 < \mathcal{O}^*(\{x_1, x_2, x_3\}, P) \approx 0.182$. This obviously proves that the measures are not ordinally equivalent. But interestingly, it also provides an argument against $\mathcal{O}'$ and in favour of $\mathcal{O}^*$. Intuitively, adding the piece of information that Tweety is a ground-dweller should not decrease the coherence of the original set. After all, this proposition is entailed by the proposition that Tweety is a penguin. But according to $\mathcal{O}'$, it does. The same argument applies for extending the set $\{x_2, x_3\}$ to $\{x_1, x_2, x_3\}$. Adding the piece of information that Tweety is a bird should not decrease the coherence of the original set, because this proposition is also entailed by the proposition that Tweety is a penguin. Irrespective of this interesting by-product, we can state the final result of this section:

Theorem 7
No pair in the set $\{\mathcal{O}, O', O^*\}$ is ordinally equivalent.

Although this result proves that there are cases for which the measures disagree, it does not tell us anything about the overall agreement between the measures. It would, however, be very interesting to see how strongly the measures agree in a highly diverse class of cases. One well-established way of addressing this issue is to analyse the measures' behaviour in a Monte Carlo simulation. Such analyses have already been provided for Bayesian confirmation measures (Tentori et al., 2007) and for extant probabilistic measures of coherence (Koscholke, 2016). For the present purpose, we generated one million uniformly distributed probability assignments over the Boolean algebra generated by three atomic propositions x_1, x_2 and x_3. For the random number generation, the Mersenne Twister algorithm (Matsumoto and Nishimura, 1998) was used. For each of these probability assignments, the corresponding values of our three measures were recorded, and for each pair of measures, Spearman's (1904) rank correlation coefficient was computed. The results are shown in Table 1, and they are not very surprising: all three measures are strongly correlated with each other—no value is below 0.9, which indicates a very high degree of statistical association. It is, however, quite interesting to see that $\mathcal{O}^*$ is more strongly correlated with both $\mathcal{O}$ and $\mathcal{O}'$ than the two are correlated with each other.

To avoid misinterpretation of these results, notice that the correlation values refer to coherence assessments for sets of *equal* size, namely, three-element sets. The situation looks very different when we compare coherence assessments for sets of different size: for instance, if for each of the simulated probability distributions we would like to know if extending a two-element set to a three-element set yields a strictly higher degree of coherence, there is no agreement at all between $\mathcal{O}$ and $\mathcal{O}^*$. This is no surprise given Theorem 1. The orderings provided by $\mathcal{O}'$ and $\mathcal{O}^*$ only agree on around 36% of the simulated distributions. In other words, the fact that two measures are highly correlated with each other does not rule out that there is some other class of cases in which they do not agree. It is the overall behaviour of $\mathcal{O}^*$ that characterizes this measure and that distinguishes it from $\mathcal{O}$ or $\mathcal{O}'$.

One interesting detail this simulation has also revealed is that although the values of $\mathcal{O}$ for a specific pair (X, P) can be lower or equal to the values of $\mathcal{O}^*$, they cannot be higher—at least in the simulated set of distributions. There might be a straightforward algebraic explanation for this. However, although we consider this an interesting observation, we do not find it interesting enough to provide a proof

Table 1 Cross-correlation matrix

	$\mathcal{O}$	$\mathcal{O}'$	$\mathcal{O}^*$
$\mathcal{O}$	1	0.928	0.986
$\mathcal{O}'$	0.928	1	0.963
$\mathcal{O}^*$	0.986	0.963	1

for this behaviour. For $\mathcal{O}'$, however, there is no such restriction: its values can be lower than, higher than, or equal to the values assigned by $\mathcal{O}^*$ to some pair (X, P). With these final remarks we close this section.

6. Conclusion

This paper has been concerned with relative overlap measures of coherence. Although the arguments provided against the plain overlap measure $\mathcal{O}$ and the refined overlap measure $\mathcal{O}'$, as discussed in §2, are indeed compelling, we have argued that this does not entail that the view that relative overlap can serve as a foundation for a proper probabilistic measure of coherence has to be given up. In fact, we have presented a new measure $\mathcal{O}^*$ in §3, which is true to the idea of coherence as relative overlap, but which we have shown to withstand the arguments. But not only this—it also satisfies a set of appealing desiderata for coherence measures, and performs flawlessly in a number of well-established test cases from the literature, as shown in §§4 and 5. Of course, this result does not mean that there are no arguments which could show that the idea of coherence as relative overlap is flawed for other reasons. But if there are such arguments, the burden of proof is on the side of those who want to argue against this idea. For us, relative overlap is back in the game.

We hope that we have been able to show that the idea of coherence as relative overlap can be re-established, and in fact opens a space of new opportunities for future research on probabilistic measures of coherence. In particular, we are very interested to see how this new measure performs in psychological investigations of coherence assessments. Studies of this kind have been carried out by Harris and Hahn (2009) and extended by Koscholke and Jekel (2015). Unfortunately, in this context we cannot provide such an investigation, and instead warmly invite other researchers to pursue this idea. Another philosophically interesting path of research is to investigate how the new measure performs with respect to truth-conduciveness. Since it has been shown by Angere (2007, 2008) in a number of computer simulations that the regular relative overlap measure $\mathcal{O}$ exhibits quite a high degree of truth-conduciveness, we are eager to see how the new measure $\mathcal{O}^*$ fares in this regard. Since the latter is based on the former and they have been shown to be highly correlated, it is to be expected that the degree of truth-conduciveness is similarly high. Still, investigations of this kind also have to be left for future research.[1]

References

Akiba, Ken 2000: 'Shogenji's Probabilistic Measure of Coherence Is Incoherent'. *Analysis*, 60(4), pp. 356–24.

Angere, Staffan 2007: 'The Defeasible Nature of Coherentist Justification'. *Synthese*, 157(3), pp. 321–35.

—— 2008: 'Coherence as a Heuristic'. *Mind*, 117, pp. 1–26.

BonJour, Laurence 1985: *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

—— 1999: 'The Dialectic of Foundationalism and Coherentism'. In John Greco and Ernest Sosa (eds.), *The Blackwell Guide to Epistemology*, pp. 117–42. Malden, MA: Blackwell Publishing.

Bovens, Luc, and Stephan Hartmann 2003a: *Bayesian Epistemology*. Oxford: Oxford University Press.

—— 2003b: 'Solving the Riddle of Coherence'. *Mind*, 112, pp. 601–33.

Bovens, Luc, and Erik J. Olsson 2000: 'Coherentism, Reliability and Bayesian Networks'. *Mind*, 109, pp. 685–719.

Crupi, Vincenzo, Katya Tentori, and Michel Gonzalez 2007: 'On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues'. *Philosophy of Science*, 74(2), pp. 229–52.

Douven, Igor, and Wouter Meijs 2007: 'Measuring Coherence'. *Synthese*, 156(3), pp. 405–25.

Fitelson, Branden 2003: 'A Probabilistic Theory of Coherence'. *Analysis*, 63(3), pp. 194–9.

[1] We would like to thank the participants of the Reasoning Club Conference 2017 at the University of Turin and the participants of the *Mind, World and Action Conference* 2017 at the Inter-University Centre Dubrovnik for helpful feedback. We are also very grateful to two anonymous reviewers whose comments helped us to improve this paper. Special thanks to Peter Brössel for taking the time to discuss the central ideas of this paper in detail. This work was funded by Grant SCHU 3080/3-1 to Moritz Schulz from the DFG as part of the Emmy Noether Group 'Knowledge and Decision' and by Grant SI 1731/1-1 to Mark Siebel from the DFG as part of the priority programme 'New Frameworks of Rationality'.

Fréchet, Maurice 1935: 'Généralisation du théorème des probabilités totales'. *Fundamenta Mathematica*, 25(1), pp. 379–87.

Glass, David H. 2002: 'Coherence, Explanation, and Bayesian Networks'. In Michael O'Neill, Richard F. E. Sutcliffe, Conor Ryan, Malachy Eaton, and Niall J. L. Griffith (eds.), *Artificial Intelligence and Cognitive Science, 13th Irish Conference, AICS 2002*, pp. 177–82. Berlin: Springer.

—— 2005: 'Problems with Priors in Probabilistic Measures of Coherence'. *Erkenntnis*, 63(3), pp. 375–85.

Harris, Adam J. L., and Ulrike Hahn 2009: 'Bayesian Rationality in Evaluating Multiple Testimonies: Incorporating the role of coherence'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(5), pp. 1366–73.

Koscholke, Jakob 2015: 'Evaluating Test Cases for Probabilistic Measures of Coherence'. *Erkenntnis*, 81(1), pp. 155–81.

—— 2016: 'Carnap's Relevance Measure as a Probabilistic Measure of Coherence'. *Erkenntnis*, 82(2), pp. 339–50.

Koscholke, Jakob and Marc Jekel 2015: 'Probabilistic Coherence Measures: A Psychological Study of Coherence Assessment'. *Synthese*, 194(4), pp. 1303–22.

Koscholke, Jakob and Michael Schippers 2015: 'Against Relative Overlap Measures of Coherence'. *Synthese*, 193(9), pp. 2805–14.

Matsumoto, Makoto, and Takuji Nishimura 1998: 'Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudorandom Number Generator'. *ACM Transactions on Modeling and Computer Simulation*, 8(1), pp. 3–30.

Meijs, Wouter 2006: 'Coherence as Generalized Logical Equivalence'. *Erkenntnis*, 64(2), pp. 231–52.

—— 2007: 'A Corrective to Bovens and Hartmann's Measure of Coherence'. *Philosophical Studies*, 133(2), pp. 151–80.

Meijs, Wouter and Igor Douven 2005: 'Bovens and Hartmann on Coherence'. *Mind*, 114, pp. 355–63.

Olsson, Erik J. 2002: 'What Is the Problem of Coherence and Truth?' *Journal of Philosophy*, 99(5), pp. 246–72.

—— 2005: *Against Coherence: Truth, Probability and Justification*. Oxford: Oxford University Press.

Rescher, Nicholas 1973: *The Coherence Theory of Truth*. Oxford: Oxford University Press.

Roche, William 2013: 'Coherence and Probability: A Probabilistic Account of Coherence'. In Michał Araszkiewicz and Jaromír Šavelka (eds.) *2013: Coherence: Insights from Philosophy,* *Jurisprudence and Artificial Intelligence*, pp. 59–91. Dordrecht: Springer.

Schippers, Michael 2014a: 'Incoherence and Inconsistency'. *Review of Symbolic Logic*, 7(3), pp. 511–28.

—— 2014b: 'Probabilistic Measures of Coherence: From Adequacy Constraints towards Pluralism'. *Synthese*, 191(16), pp. 3821–45.

Schippers, Michael and Mark Siebel 2015: 'Inconsistency as a Touchstone for Coherence Measures'. *Theoria*, 30(1), pp. 11–41.

Schupbach, Jonah N. 2011: 'New Hope for Shogenji's Coherence Measure'. *British Journal for the Philosophy of Science*, 62(1), pp. 125–42.

Shogenji, Tomoji 1999: 'Is Coherence Truth Conducive?' *Analysis*, 59(4), pp. 338–45.

—— 2001: 'Reply to Akiba on the Probabilistic Measure of Coherence'. *Analysis*, 61(2), pp. 147–50.

Siebel, Mark 2004: 'On Fitelson's Measure of Coherence'. *Analysis*, 64(2), pp. 189–90.

Siebel, Mark and Werner Wolff 2008: 'Equivalent Testimonies as a Touchstone of Coherence Measures'. *Synthese*, 161(2), pp. 167–82.

Spearman, Charles 1904: 'The Proof and Measurement of Association between Two Things'. *American Journal of Psychology*, 15(1), pp. 72–101.

Tentori, Katya, Vincenzo Crupi, Nicolao Bonini, and Daniel Osherson 2007: 'Comparison of Confirmation Measures'. *Cognition*, 103, pp. 107–19.

