Coherence and Common Causes: Against Relevance-sensitive Measures of Coherence

Jakob Koscholke and Michael Schippers

Abstract

Changing weather conditions and barometer changes usually coincide. Accordingly, the propositions that my barometer falls and that the weather conditions deteriorate are quite coherent – especially under the assumption that there is a drop in atmospheric pressure. Nevertheless, scenarios like this involving common causes turn out to be highly problematic for a prominent class of probabilistic coherence measures, namely those explicating coherence based on the idea of *relevance-sensitivity*. In this paper we show that none of these measures accords with the intuition that in the light of a common cause, two propositions referring to the cause's effects should turn out coherent. This result casts doubts on the view that these measures can be considered proper explications of the concept of coherence.

1 Introduction

Recent years have seen a considerable number of formal approaches to explicating the concept of coherence by means of probability theory, resulting in so-called *probabilistic* *coherence measures*. It is an enduring task to evaluate these measures with respect to their adequacy. This paper contributes to this project by concentrating on a well-known family of measures that are based on the idea that the concept of coherence is best explicated in terms of probabilistic relevance (Shogenji [1999]; Fitelson [2003]; Douven and Meijs [2007]; Schupbach [2011]). These measures will be evaluated in a hitherto neglected class of test cases involving a pair of events that share a common cause.

Common causes are ubiquitous in science as well as in everyday life. Typical examples are changes in atmospheric pressure causing both barometer and weather changes or smoking causing both yellow fingers and lung cancer. Causal structures like these are also a common topic of philosophical theories of causation. As regards *regularity* theories of causation (Hume [[1748] 1999]; Mackie [1965]) common cause structures often given rise to empty regularities posing a challenge for these theories. As regards *probabilistic* theories of causation (Reichenbach [1956]; Suppes [1970]) common causes clearly reveal the need for more sophistication than just the equation of causal and probabilistic relevance.

As we will show in this paper, common cause structures also pose a threat for probabilistic theories of *coherence*. To illustrate this, let x be the proposition that my barometer falls and y the proposition that the weather conditions deteriorate. Then the set X containing x and y as its members seems quite coherent given that usually my barometer drops when the weather conditions change and vice versa. Moreover, in the light of a further piece of information z that there is a drop in atmospheric pressure, the set X still seems quite coherent if not even more coherent because the event expressed in z is a common cause of both phenomena. Nevertheless, all coherence measures based on the idea of *relevance-sensitivity* (to be explained below) assess the set X as neither coherent nor incoherent in the light of z, independently of what X's degree of coherence is irrespective of z. This, we conclude, speaks against the adequacy of this class of measures.

The paper is structured as follows: in section 2 we introduce the family of relevance-sensitive coherence measures. In section 3 we show that all members of this family do not adequately handle test cases featuring common causes while in section 4 we discuss potential objections against this result. Finally, we summarize our conclusions in section 5.

2 Relevance-sensitive Coherence Measures

Relevance-sensitive measures of coherence are based on Bayesian measures of evidential support, sometimes also referred to as (incremental) confirmation. Both kinds of measures presuppose the following formal apparatus: let A be an algebra of propositions and let $P : A \to [0,1]$ be a probability measure over A, i.e., a non-negative, real-valued function such that $P(x) = 1$ if $x \in A$ is a tautology, and P is finitely additive so that $P(x \vee y) = P(x) + P(y)$ if $x \in A$ and $y \in A$ are logically incompatible. Conditional probabilities are defined as usual by $P(x|y) = P(x \wedge y)/P(y)$ for any $y \in A$ with $P(y) > 0$.

Furthermore, let $\mathbf{P}$ denote the set of all probability functions P over A. Then a Bayesian measure of support can be characterized as a function $S : A^3 \times \mathbf{P} \to \mathbb{R}$ where the real number assigned by S to the quadruple (x, y, z, P) is supposed to represent the degree to which a hypothesis x is supported or disconfirmed by a piece of evidence y in the light of background knowledge z under some joint probability assignment P. Over the years, a variety of such measures has been proposed and discussed in the literature (Crupi *et al.* [2007]; Festa [2012]). Table 1 gives an overview over some prominent measures (ordinal equivalent versions are omitted) and their proponents.

Measure	**Proponent**
$P(x\|y \wedge z)/P(x\|z)$	Keynes [1921]
$P(x\|y \wedge z) - P(x\|z)$	Carnap [1950]
$P(x \wedge y \wedge z) \times P(z) - P(x \wedge z) \times P(y \wedge z)$	Carnap [1950]
$[P(y\|x \wedge z) - P(y\|\neg x \wedge z)]/[P(y\|x \wedge z) + P(y\|\neg x \wedge z)]$	Kemeny and Oppenheim [1952]
$[P(x\|y \wedge z) - P(x\|z)]/[1 - P(x\|z)] \times P(y\|z)$	Rescher [1958]
$P(y\|x \wedge z) - P(y\|\neg x \wedge z)$	Nozick [1981]
$P(y\|x \wedge z) - P(y\|z)$	Mortimer [1988]
$P(x\|y \wedge z) - P(x\|\neg y \wedge z)$	Christensen [1999]
$\min\{P(x\|y \wedge z), P(x\|z)\}/P(x\|z) - \min\{P(\neg x\|y \wedge z), P(\neg x\|z)\}/P(\neg x\|z)$	Crupi *et al.* [2007]
$[\log_2 P(x\|y \wedge z) - \log_2 P(x\|z)]/ - \log_2 P(x\|z)$	Shogenji [2012]

Table 1: Bayesian measures of evidential support.

Notice that usually the simplified, unconditional versions of these measures are discussed in the literature. Quite obviously, those versions can be obtained from the conditional version presented here when assuming the background knowledge z to be a tautology $\top$. In this context, however, we obviously need to depart from this restriction and employ the conditionalized versions since the assumption of non-tautological but contingent background knowledge is going to play a key role in our argument. Furthermore notice that the ordinal differences between the various confirmation measures will not matter in the following. The important point of all these functions is that they are *relevance-sensitive* in the sense that they feature a neutrality threshold τ separating positive and negative degrees of support along the following lines:

(†) $S(x, y, z, P) > / = / < \tau$ if and only if $P(x|y \wedge z) > / = / < P(x|z)$.

What (†) tells us is that the degree of support assigned by S differs from the measures' neutral value τ if and only if x's probability given y *and* z differs from x's probability given only z. This condition is usually assumed to be the foundation of any support measure labelled as *Bayesian* (Crupi *et al.* [2007]).

Let us now turn to probabilistic coherence measures. In addition to the conventions introduced above, let $2^A_{\geq 2}$ denote the powerset of A without the empty set and all singleton sets.[1] Then a probabilistic coherence measure is a function $C : 2^A_{\geq 2} \times A \times \mathbf{P} \to \mathbb{R}$

[1]The exclusion of empty and singleton sets in this context is known as "Rescher's principle" Olsson [2005, 17] according to which coherence can only meaningfully assigned to sets involving at least two propositions.

assigning triples (X,z,P) real numbers that are supposed to represent the degree of coherence of X in the light of z under probability distribution P.[2] In what follows we limit our investigations to cases in which the algebra A is based on three propositions x,y,z under some probability function P so that we would like to assess the degree of coherence of the set $X = \{x,y\}$ in the light of z under P. Now, analogously to the aforementioned condition (†) for Bayesian support measures, a probabilistic coherence measure is *relevance-sensitive* if there is a threshold τ' separating coherent and incoherent sets of propositions such that the following condition is satisfied:

$$(\dagger') \;\; C(X,z,P) > / = / < \tau' \text{ if and only if } P(x \wedge y|z) > / = / < P(x|z) \times P(y|z)$$

According to $(\dagger')$ the degree of coherence assigned to X in the light of z by a relevance-sensitive coherence measure C deviates from the threshold τ' if and only if x and y are not probabilistically independent given z. On the other hand, measures satisfying $(\dagger')$ assign a neutral degree of coherence to the set X if and only if X's elements x and y are independent given z.

Over the years a large number of relevance-sensitive coherence measures has been proposed in the literature. The starting point is usually assumed to be (Shogenji [1999]) where Shogenji suggests to quantify coherence by means of the fraction that takes the probability of the conjunction over all of the set's members as its nominator, and the product of the marginal probabilities of all of the set's members as its denominator. Schupbach has renewed the interest in this measure in (Schupbach [2011]) by endorsing a version of Shogenji's measure that is sensitive to the coherence of subsets of the set under consideration. However, with respect to our test cases featuring only pairs of propositions, the measures are easily seen to be ordinally equivalent, i.e. their coherence-based orderings of two-membered sets of propositions do not differ. What is more, in cases involving two-membered sets of propositions Shogenji's measure also turns out to be ordinally equivalent to a measure that will be introduced below. Accordingly, we will not consider Shogenji's and Schupbach's measures separately. Instead, we keep in mind that their measures are also affected by the problem discussed in this paper. Similarly, Fitelson's measure and the refined version he has proposed later (Fitelson [2003], [2004]) are also special cases of the following recipe that has been developed in (Douven and Meijs [2007]). This recipe takes the idea of coherence as average mutual support as its starting point and allows utilizing a Bayesian support measure S to quantify coherence. Now, let $X = \{x,y\}$ be a set of propositions under

[2]It is worth mentioning that in many formal representations of probabilistic coherence measures the background assumption z is usually omitted – for exceptions see Wheeler [2009] and Schippers [2015]. This omission, however, is equivalent to assuming that z is a tautology $\top$. In this sense, the notion of an *unconditional* degree of coherence is just a special case of the notion of a *conditional* degree of coherence. And since there might be cases in which one would like to assess the degree of coherence in the light of some contingent proposition – e.g. in order to compare the degree of coherence of some set X against different background assumptions or their negations – the restriction to tautologous background assumptions should be relaxed. Moreover, given that *confirmation* is usually modelled as a four-place relation that explicitly takes into account the background knowledge, we see no reason for neglecting this assumption when proceeding from confirmation to coherence.

some joint probability function P, then X's degree of coherence in the light of z is equated with the average degree of mutual support between x and y in the light of z:

$$C_S(X,z,P) = 1/2 \times [S(x,y,z,P) + S(y,x,z,P)] \tag{1}$$

Thus, if x supports y in the light of z under P and vice versa, then the set X is coherent. On the other hand, X is neither coherent nor incoherent if and only if neither x supports y in the light of z under P nor does y support x in the light of z under P. The following observation is a straightforward consequence:

Observation 1 *If S satisfies* $(\dagger)$, *then C_S satisfies* $(\dagger')$.

In other words, every Bayesian support measure that is relevance-sensitive in the sense of $(\dagger)$ yields, when plugged into Douven and Meijs' recipe (1), a coherence measure that is relevance-sensitive in the sense of $(\dagger')$. This means that the threshold values τ of S and τ' of C_S will be identical. Given this characterization, we now turn to our evaluation of test cases involving common causes.

3 Common causes, Screening-off and Coherence

A common cause structure is given by a setting in which two events x and y are caused by the same event z.[3] A graphical representation of such a structure is given in Figure 1 below:

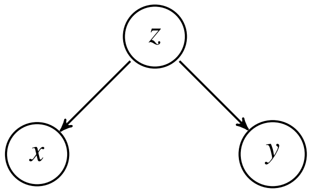

Figure 1: Common cause structure. Arrows indicate causal relevance relations.

Causal structures like these pose a challenge for philosophical theories of causation equating probabilistic and causal relevance. In the context of such theories common cause scenarios are usually modelled by the following (naive) assumptions:

(i) $P(x|z) > P(x)$ and $P(y|z) > P(y)$

(ii) $P(x \wedge y) > P(x) \times P(y)$

(iii) $P(x \wedge y|z) = P(x|z) \times P(y|z)$

[3]Here and in what follows we will use x, y and z *both* to denote events that can be the relata of the causal relevance relation *and* the corresponding propositions stating that the event is instantiated.

While condition (i) represents the causal relevance the common cause has for its two effects, the second assumption guarantees that the two effects are positively correlated with each other. However, it is obvious that these assumptions are not sufficient. For example, although the fact that my barometer falls (x) is *probabilistically* relevant to the fact that the weather conditions deteriorate (y) – formally captured by condition (ii) or equivalently by $P(y|x) > P(y)$ or $P(x|y) > P(x)$ – it is obviously not *causally* relevant. Therefore, probabilistic theories of causation are typically assumed to require condition (iii) known as *screening-off* (Reichenbach [1956]).[4] An event z screens-off two other events x and y from each other if and only if y and x are independent in the light of z – formally captured by condition (iii) or equivalently by $P(x|y \wedge z) = P(x|z)$ or $P(y|x \wedge z) = P(y|z)$. In this sense, the fact that there is a drop in atmospheric pressure (z) screens-off deteriorating weather conditions (y) and my falling barometer (x) from each other and hence x cannot be considered as a proper cause for y just as y cannot be considered as a proper cause for x.

Now, how does the situation look with respect to coherence? This question is especially interesting since the impact of causal relations on coherence assessments has been largely neglected by probabilistic coherence theorists. Intuitively, the propositions that my barometer falls (x) and that the weather conditions deteriorate (y) are quite coherent in the absence of any background assumption, i.e. in the light of $\top$. This intuition is driven by the fact that the events to which the propositions refer usually coincide – as modelled by the correlation assumption (ii). In addition to this intuition, the two propositions remain equally coherent or might even get more coherent if instead of assuming $\top$ one assumes that there is a drop in atmospheric pressure (z). This additional intuition is based on the fact that z is the common cause of x and y – as modelled by assumption (i). Thus, one is provided information about the nature of the association between x and y and the reason for it: they are not just correlated, but they are correlated because they are both effects of the same cause z. Hence, there seems to be no good reason why $\{x,y\}$ should be less coherent in the light of their common cause z than in the light of a tautology $\top$. The fact that the correlation between x and y vanishes under the assumption of z should not generally mean that their coherence does. This view, however, is violated by all relevance-sensitive coherence measures considered above, as is captured by the following observation:

Observation 2 *For any joint probability function P over propositions x,y,z: if x is screened-off from y by z, then $C(\{x,y\},z,P) = \tau'$ for all measures C that satisfy $(\dagger')$.*

This observation immediately entails the following corollary:

[4] Note that the screening-off condition that is used in this paper is faced with some problems; among others, a serious drawback is that according to this condition causation is *not* transitive in causal chains, i.e. in a situation with three events x,y and z, where x is causally relevant to y which in turn is causally relevant to z, x cannot also be considered causally relevant to z because its probabilistic impact on z is screened-off by y. There are refined screening-off conditions that avoid this problem (Suppes [1970]). However, the focus of the present paper are coherence relations obtaining in common cause structures, so that we will dispense with discussions of the pros and cons of alternative screening-off conditions.

Corollary 1 *All coherence measures in line with recipe* (1) *based on support measures satisfying* (†) *assess two propositions x, y referring to two effects as neither coherent nor incoherent in the light of their corresponding common cause z.*

This result is problematic in two ways. First, there seems to be no good reason to judge any set of propositions referring to two correlated effects x, y as neutral with respect to its coherence in the light of their common cause z. In fact, it seems more adequate to judge this set of propositions as quite coherent as we have argued. Second, it is problematic because it is possible that relevance-sensitive coherence measures judge a set of propositions as coherent in the absence of any background assumption, i.e. z is taken to be a tautology $\top$, but neutrally and therefore less coherent in the light of the common cause z. Intuitively, however, two propositions referring to two effects of a common cause are already coherent and should only be equally or more coherent in the light of the common cause than in the absence of any background assumption.

To further illustrate this point one can think of a historical scenario involving patients with two kinds of symptoms that often coincide. Let x be the proposition that the patient has symptom X and y be the proposition that she has symptom Y. Furthermore, let P be the probability distribution that captures their coincidence. Intuitively, we think that anyone would assess the propositions x and y coherent even in the absence of any further pieces of background information simply because we find both symptoms often in common. Now, assume that some time later scientists discover an unknown virus that both causes symptom X and symptom Y. Moreover, let P' be the probability distribution that results from conditionalizing on this new piece of information. Here it seems to us that although x and y are now P'-independent, they intuitively remain coherent to a degree which not lower than under P – because we know that both the corresponding symptoms are the effects of a common cause.

Hence, the interpretation of our result is quite simple: *no* relevance-sensitive coherence measure adequately handles scenarios involving common cause structures since these scenarios usually presuppose conditions like screening-off. This is highly problematic since common cause structures are by no means exotic cases. Quite the contrary, they are very common in both science and everyday contexts. Therefore, we conclude that relevance-sensitive measures cannot be considered proper explications of the concept of coherence.

4 Discussion

We argued on the basis of a general scheme of cases involving common causes that all relevance-sensitive coherence measures should be considered inadequate explications of the concept of coherence. The result that substantiates this claim is much stronger than usual test-case results employed in this context (for a recent survey see Koscholke [2016]). While those cases usually only specify one scenario as a touchstone for coherence measures, the general scheme employed here forms a whole class of test cases such that the result holds for *every* probability distribution satisfying the conditions described above. Nevertheless, one might have objections against the points made here. The two most obvious objections will be discussed in the following two subsections.

4.1 Multiple Explications of Coherence

Critics might argue that the result presented here is not strong enough to allow for reasonable judgments regarding the adequacy of certain classes of probabilistic coherence measures. Perhaps, so the argument might go, we are in a situation similar to the one alluded to by Carnap in his discussion on the concept of confirmation:

> It has repeatedly occurred in the history of science that a vehement but futile controversy arose between the proponents of two or more *explicata* who shared the erroneous belief that they had the same explanandum; when finally it became clear that they meant different explicanda, unfortunately designated by the same term, and that the different explicata were hence compatible and moreover were found to be equally fruitful scientific concepts, the controversy evaporated into nothing.
> (Carnap [1950], 26)

Might it not be the case that there are as well two or more pretheoretic concepts of coherence that have to be kept apart carefully in order to forestall confusions? Might it not be the case that what we have shown is that there is a class of test cases, i.e. cases involving common cause structures, that are coherent in a sense that is not properly captured by relevance-sensitive measures of coherence?

Let us start with the first question, i.e. whether there is evidence regarding the existence of multiple pretheoretic concepts of coherence? Indeed, we think that evidence to this effect abounds. Although the vast majority of measures discussed in the literature (for instance Olsson and Schubert [2007]; Siebel and Wolff [2008]) is based on relevance-sensitivity intuitions, there are some exceptions: these are coherence measures based on relative set-theoretic overlap as proposed in (Glass [2002]; Olsson [2002]) and in a refined manner in (Meijs [2006]), and a measure (discussed below) based on Douven and Meijs' approach and the conditional probability (Roche [2013]). As was recently shown in (Koscholke and Schippers [2015]), the former class of measures suffers from a serious drawback: according to the Glass-Olsson overlap measure of coherence it is impossible to increase coherence by expanding a set of propositions. More precisely, let X be a set of propositions under some probability distribution P, then the degree of coherence assigned to X under P by the Glass-Olsson measure will always be at least as high as the degree of coherence assigned to $X \cup \{x\}$ under P for any arbitrary proposition x. This is clearly a devastating property for a coherence measure. Just think of a set X containing the two propositions 'Tweety is a bird' and 'Tweety cannot fly', then it seems plain as day that adding the proposition 'Tweety is a penguin' to this set of propositions increases its degree of coherence (considerably) (Bovens and Hartmann [2005]). As was also shown in (Koscholke and Schippers [2015]), this property has immediate consequences for Meijs' subset-sensitive overlap measure based on the Glass-Olsson measure. According to Meijs' refined overlap measure of coherence, the degree of coherence of any set of propositions cannot be higher than its most-coherent, two-element subset. More precisely, let X be a set of propositions under some probability distribution P, then no expansion X^* of X under P can be more coherent than the most coherent subset $X' \subset X^*$ with $|X'| = 2$ (and this result holds irrespective of the chosen weighting system to aggregate the various coherence degrees of different subsets). Hence, it can be concluded that overlap measures cannot be considered proper measures of coherence.

Another extant coherence measure is based on a different concept of confirmation, sometimes referred to as *absolute confirmation* or *firmness*: a piece of evidence y is said to confirm a hypothesis x in the absolute sense if and only if x is more probable in the light of y than its negation is. More precisely, we get the following condition:

(‡) $S(x,y,z,P) > / = / < \tau$ if and only if $P(x|y \wedge z) > / = / < P(\neg x|y \wedge z)$.

Hence, absolute confirmation is only sensitive to the conditional probabilities of x and $\neg x$ given the possibly confirmatory evidence y. Roche's ([2013]) coherence measure now simply replaces the relevance-sensitive measure of (incremental) confirmation S that is commonly used when applying Douven and Meijs' recipe (1) by a measure of absolute confirmation A. Thus, we get the following measure:

$$C_A(X,z,P) = 1/2 \times [P(x|y \wedge z) + P(y|x \wedge z)]$$

Now, in the common cause scenarios discussed above both conditional probabilities involved in computing the degree of coherence of the set $X = \{x,y\}$ according to C_A are close to their maximum probability 1. Hence, C_A judges this set highly (if not maximally) coherent in the light of the common cause z.

Now let us turn to the second question considered above, i.e. whether it might be the case that scenarios involving common cause structures allude to an intuition that is more in line with the absolute confirmation based measure of coherence. We do not think that this is the case. However, the classification of test cases with respect to different coherence intuitions is often a difficult endeavour. Imagine a situation in which a pair of propositions is intuitively coherent but fails to be judged so on one of the major probabilistic approaches to coherence. Should this be considered to prove the deficiency of this latter approach to coherence? Or should we rather conclude that what is shown is that the test case addresses a coherence intuition that is not captured by this approach?

There seem to be clear-cut cases where the latter answer seems the correct one. For example, consider the following scenario borrowed from (Siebel [2004]): there are 10 independent and equally likely suspects for a murder; 8 suspects committed a robbery, 8 suspects committed a pickpocketing and 6 committed both crimes. Now consider the following two propositions:

p_1: The murderer committed a robbery.

p_2: The murderer committed a pickpocketing.

Obviously, p_1 and p_2 are subcontrary propositions, i.e. they cannot be false together. It is easy to show for subcontrary propositions that each proposition's conditional probability given the other falls short of its unconditional probability (Siebel [2004]). Hence, p_1 and p_2 will be assessed incoherent by all relevance-sensitive coherence measures (Koscholke [2016]). What is at play here seems to be an intuition that is more akin to the absolute confirmation based approach to coherence.

Perhaps we can conclude the same as regards common causes? It does not seem so. What is at issue in the specified class of scenarios is the probabilistic *correlation* between two events x and y that is explained by the presence of a common cause z and this correlation is most naturally captured by relevance-sensitive measures of coherence. Furthermore, although it is the case that in our barometer case both conditional probabilities $P(x|y)$ and $P(y|x)$ are high, nothing prevents us from considering different examples where these conditional probabilities are rather low (or fall short of .5). Given that we face a situation, where x and y are nonetheless probabilistically correlated and thus coherent they seem to remain coherent even if this probabilistic correlation vanishes under the assumption of a common cause.

4.2 Conditionalization versus Set Expansion

Another reason for concern might relate to the employed conditionalization-based approach to assessing coherence. Why, one might ask, should we opt for a model that compares the degrees of coherence of the set X containing the two effects x and y against two different backgrounds, namely *tautologous* versus *common cause* rather than expanding the set X by z? More formally, why should we compare $C(X,\top,P)$ and $C(X,z,P)$ instead of the more common comparison between $C(X,\top,P)$ and $C(X\cup\{z\},\top,P)$?

There are at least two reasons why the conditionalization-based model seems preferable to a set-theoretic expansion model. First, recall our phrasing of the relevant question above: is it true for relevance-based measures of coherence that the effect-set X is no less coherent *in the light of* its common cause z than without this latter assumption? It seems that the most natural model of this situation is by means of *learning* that z is the case, i.e. that the common cause is present. In Bayesian terms, this amounts to *conditionalizing* on the new information z. Second, as regards the expansion-approach to modelling common cause scenarios, even in fortunate settings it is possible that the degree of coherence of the expanded set falls below the degree of coherence of the reduced set. What is meant by fortunate settings? When searching for a probability model involving two effects of a common cause, it seems that more needs to be fixed than just the assumptions (i)-(iii). Think of a probability distribution in which these conditions hold but $P(x|y\wedge z)\ll P(x|y)$ so that one (or both) of the effects becomes (considerably) less probable in the light of the common cause. Such a probability distribution does not seem true to the idea of a common cause scenario.

What is more, the common cause z should be probabilistically relevant for its effect, i.e. the conditional probabilities of x and y conditional on their common cause z should exceed their respective unconditional probabilities. Accordingly, in addition to the conditions (i)-(iii) we assume that an exemplary probability model for a common cause scenario satisfies the following condition:

(iv) $P(x|y \wedge z) \geq P(x|y)$ and $P(y|x \wedge z) \geq P(y|x)$

It can be shown that even under these conditions it is possible that the degree of coherence assigned to the expanded set $\{x,y,z\}$ *undercuts* the degree of coherence of the reduced set $\{x,y\}$. Simply consider the following probability distribution:

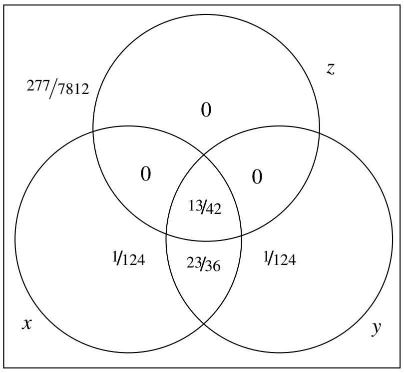

Figure 2: Probability distribution for a common cause scenario.

It can easily be seen that this distribution satisfies the conditions (i)-(iv) mentioned beforehand. However, as Table 2 indicates, the vast majority of considered measures (all but the measure based on Keynes' relevance quotient) agree in that the larger set $\{x,y,z\}$ is *less* coherent than the reduced set $\{x,y\}$ containing only the propositions referring to the effects.[5]

[5]Notice that the distribution given in Figure 2 has been found using a custom written program in GNU Octave. Unfortunately, we have not been able to find a counterexample for the coherence measure based on Keynes' ratio measure of confirmation, not even by using Branden Fitelson's MATHEMATICA application **PrSAT** (Fitelson [2008]). Neither have we been able to find a proof showing that there is no such counterexample. Accordingly, one might be inclined to consider this result a partial vindication of the ratio-based measure of coherence. This, however, seems problematic in at least two respects. On the one hand, it seems perfectly reasonable to assume that there *is* such a counterexample but we simply have not been able to find it. On the other hand, there are numerous results clearly indicating the deficiency of this coherence measure (Koscholke [2016]; Schippers and Siebel [2015]). Finally, we have been able to find numerous distributions for the coherence measure based on Keynes measure once we dropped assumption (iv) above.

Measure	$\{x,y\}$	$\{x,y,z\}$
Keynes	1.037	1.046
Carnap (difference measure)	0.035	0.031
Carnap (relevance measure)	0.034	0.017
Kemeny and Oppenheim	0.685	0.545
Rescher	0.771	0.266
Nozick	0.806	0.297
Mortimer	0.035	0.031
Christensen	0.806	0.297
Crupi et al.	0.806	0.560
Shogenji	0.810	0.568
Roche	0.992	0.717

Table 2: Results for the distribution shown in Figure 1.

Note that this does seem a remarkable fact regarding the adequacy of relevance-sensitive coherence measures on its own: when BonJour set out his coherence theory of justification, one of the key conditions to boost coherence he mentions is the presence of inferential relations obtaining among the set of beliefs under consideration (BonJour [1985]). Now the given probability distribution 2 illustrates that contrary to intuition we can have a situation in which we *add* a proposition providing (deductive) inferential relations to a set of propositions and thereby *reduce* the coherence of the set. This result, too, casts serious doubts on the tenability of a large class of relevance-based coherence measures.

Hence, whichever way we look at it, i.e. whether we prefer the conditionalization-based approach or the set-theoretic expansion approach, relevance-sensitive coherence measures simply provide inadequate coherence assessments in common cause scenarios as described before. By contrast, this does not hold for coherence measures based on the recipe (1) by Douven and Meijs and an absolute notion of support. The conditions (i)-(iv) together with the conditionalization strategy for modelling coherence assessment in common cause scenarios entail that $C_A(X,z,P) \geq C_A(X,\top,P)$.

5 Conclusion

The results discussed above clearly indicate that relevance-sensitive coherence measures fail in an important class of cases, namely common cause scenarios. No matter how we model the coherence assessment in such cases, be it conditionalization-based or as set-theoretic expansion, the measures simply yield counter-intuitive results. We also argued that the failure of relevance-sensitive coherence measures is not due to unsuitable coherence intuitions underlying the specified class of cases. Quite the contrary, common cause scenarios are prime examples of coherence intuitions based on correlations and should therefore be mastered by relevance-sensitive measures of coherence. The fact that these measures do not does not indicate a different coherence intuition. Rather, it indicates that these measures do not capture the concept of coherence they are supposed to capture. We therefore conclude that these measures are inadequate.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program SPP 1516 New Frameworks of Rationality (grant SI 1731/1-1 to Mark Siebel). We are thankful to two anonymous referees for providing comments that helped us improving this paper.

University of Oldenburg
Philosophy Department
26111 Oldenburg, Germany
jakob.koscholke@uni-oldenburg.de

References

BonJour, L. [1985]: *The Structure of Empirical Knowledge*, Cambridge: Harvard University Press.

Bovens, L. and Hartmann, S. [2005]: 'Why there cannot be a single probabilistic measure of coherence', *Erkenntnis*, **63**, pp. 361–374.

Carnap, R. [1950]: *Logical foundations of probability*, Chicago: University of Chicago Press.

Christensen, D. [1999]: 'Measuring Confirmation', *Journal of Philosophy*, **96**, pp. 437–461.

Crupi, V., Tentori, K. and Gonzales, M. [2007]: 'On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues', *Philosophy of Science*, **74**, pp. 229–252.

Douven, I. and Meijs, W. [2007]: 'Measuring Coherence', *Synthese*, **156**, pp. 405–425.

Festa, R. [2012]: '"'For Unto Every One That Hath Shall Be Given"'. Matthew Properties for Incremental Confirmation', *Synthese*, **184**, pp. 89–100.

Fitelson, B. [2003]: 'A Probabilistic Theory of Coherence', *Analysis*, **63**, pp. 194–199.

Fitelson, B. [2004]: 'Two technical corrections to my coherence measure', `http://fitelson.org/coherence2.pdf`.

Fitelson, B. [2008]: 'A Decision Procedure for Probability Calculus with Applications', *The Review of Symbolic Logic*, **1**, pp. 111–125.

Glass, D. H. [2002]: ‘Coherence, Explanation, and Bayesian Networks’, in M. O’Neill, R. F. E. Sutcliffe, C. Ryan, M. Eaton and N. J. L. Griffith (*eds*), *Artificial Intelligence and Cognitive Science. 13th Irish Conference, AICS 2002, Limerick, Ireland, September 2002*, Berlin: Springer, pp. 177–182.

Hume, D. [[1748] 1999]: *An Enquiry Concerning Human Understanding*, Oxford: Oxford University Press.

Kemeny, J. and Oppenheim, P. [1952]: ‘Degrees of Factual Support’, *Philosophy of Science*, **1952**, pp. 307–324.

Keynes, J. [1921]: *A Treatise on Probability*, London: Macmillan.

Koscholke, J. [2016]: ‘Evaluating Test Cases for Probabilistic Measures of Coherence’, *Erkenntnis*, **81**(1), pp. 155–181 DOI: 10.1007/s10670-015-9734-1.

Koscholke, J. and Schippers, M. [2015]: ‘Against relative overlap measures of coherence’, *Forthcoming in Synthese*, pp. 1–10 DOI: 10.1007/s11229-015-0887-x.

Mackie, J. L. [1965]: ‘Causes and conditions’, *American Philosophical Quarterly*, **2**, pp. 245–264.

Meijs, W. [2006]: ‘Coherence as Generalized Logical Equivalence’, *Erkenntnis*, **64**, pp. 231–252.

Mortimer, H. [1988]: *The Logic of Induction*, Paramus: Prentice Hall.

Nozick, R. [1981]: *Philosophical Explanations*, Oxford: Clarendon.

Olsson, E. J. [2002]: ‘What is the Problem of Coherence and Truth?’, *The Journal of Philosophy*, **94**, pp. 246–272.

Olsson, E. J. [2005]: *Against Coherence: Truth, Probability and Justification*, Oxford: Oxford University Press.

Olsson, E. J. and Schubert, S. [2007]: ‘Reliability conducive measures of coherence’, *Synthese*, **157**(3), pp. 297–308.

Reichenbach, H. [1956]: *The Direction of Time*, Berkeley: University of California Press.

Rescher, N. [1958]: ‘Theory of evidence’, *Philosophy of Science*, **25**, pp. 83–94.

Roche, W. [2013]: ‘Coherence and probability: A probabilistic account of coherence’, in M. Araszkiewicz and J. Savelka (*eds*), *Coherence: Insights from philosophy, jurisprudence and artificial intelligence*, Dordrecht: Springer, pp. 59–91.

Schippers, M. [2015]: 'The Problem of Coherence and Truth Redux', *Erkenntnis*, pp. 1–35.

Schippers, M. and Siebel, M. [2015]: 'Inconsistency as a Touchstone for Coherence Measures', *Theoria*, **30**, pp. 11–41.

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

Shogenji, T. [1999]: 'Is Coherence Truth Conducive?', *Analysis*, **59**, pp. 338–345.

Shogenji, T. [2012]: 'The degree of epistemic justification and the conjunction fallacy', *Synthese*, **184**, pp. 29–48.

Siebel, M. [2004]: 'On Fitelson's measure of coherence', *Analysis*, **64**, pp. 189–190.

Siebel, M. and Wolff, W. [2008]: 'Equivalent testimonies as a touchstone of coherence measures', *Synthese*, **161**, pp. 167–182.

Suppes, P. [1970]: *A Probabilistic Theory of Causality.*, Amsterdam: North Holland.

Wheeler, G. [2009]: 'Focused Correlation and Confirmation', *British Journal for the Philosophy of Science*, **60**(1), pp. 79–100.