



# LUND UNIVERSITY

## What Is the Problem of Coherence and Truth?

Olsson, Erik J

*Published in:*  
Journal of Philosophy

2002

[Link to publication](#)

*Citation for published version (APA):*

Olsson, E. J. (2002). What Is the Problem of Coherence and Truth? *Journal of Philosophy*, XCIX(5), 246-272.

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00



# WHAT IS THE PROBLEM OF COHERENCE AND TRUTH?\*

Erik J. Olsson

Olsson, E. J. (2002). What Is the Problem of Coherence and Truth?. *The Journal of Philosophy*, XCIX, 246-272.

## I. INTRODUCTION

In inquiry we often proceed by gathering information from different sources that may not be very trustworthy in themselves, and all we have to go on may well be the degree to which the reports cohere. If they do cohere or agree to a large extent we tend to think that the information is credible. For instance, if the first dubious witness to be queried says that John was at the crime scene, the second that John has a gun and the third that John shortly after the robbery transferred a large sum to his bank account, then the striking coherence of the different testimonies would normally make us pretty confident that John is to be held responsible for the act, their individual dubiousness notwithstanding.

Impressed with this sort of examples, some philosophers have been led to the conclusion that coherence is the one single thing separating warranted from unwarranted beliefs. Roughly speaking, beliefs that hang well together, agree or exhibit mutual support are, according to philosophers of this inclination, thereby justifiably held and, if true, known to be so. Others have been more modest in their philosophical extrapolations from such pieces of common sense reasoning. For C. I. Lewis, for example, coherence cannot create credibility from scratch; it can only amplify an already existing positive degree thereof. Yet coherence has an important role to play also in Lewis's foundationalist framework in which it is needed for building up probabilities "sufficient for rational and practical reliance."<sup>1</sup>

Lewis, as we just saw, can be interpreted as suggesting a probabilistic rendering of what it means for coherence to imply truth. But what exactly is the role of coherence, and what are the probabilistic facts here? In their search for answers, a number of recent authors have focused on

---

\* Thanks are due to Ludwig Fahrback for patiently going through several earlier versions of this paper with me. I have also benefited from the criticism and suggestions of Hans Rott, Tomoji Shogenji and two anonymous referees of this journal. My research was financed by the DFG (*Deutsche Forschungsgemeinschaft*) as a contribution to the project *Logik in der Philosophie*.

<sup>1</sup> *An Analysis of Knowledge and Valuation* (La Salle: Open Court, 1946), p. 357. Lewis sees himself rather as advocating what he calls "conceptual pragmatism," a philosophical position situated "somewhere between [William] James and the absolute pragmatism of [Josiah] Royce" (p. 17 in his "Autobiography," in P. A. Schilpp, ed., *The Philosophy of C. I. Lewis*, La Salle: Open Court, 1968). All references to Lewis's work in the following concern his 1946 book.

the problem of whether more coherence implies a higher likelihood of truth.<sup>2</sup> Let us say that coherence is *truth conducive* if it has that property. The question then is: If a system *S* is more coherent than another system *S'*, are we then allowed to conclude that *S* is more likely than *S'* to be true as a whole?

There are good reasons to pay attention to this particular question. First, as Peter Klein and Ted A. Warfield<sup>3</sup> point out, it asks for a minimal sense in which coherence could imply truth. It would seem difficult to maintain that coherence implies truth without also maintaining that more coherence implies a higher likelihood of truth. Second, it is relatively clear and unambiguous. By contrast, the question “Is a coherent system highly likely to be true?”, for instance, suffers from serious vagueness—indeed, along two dimensions: concerning how high a degree of coherence it takes for a system to qualify as “coherent” as well as concerning how high a degree of likelihood it takes for a system to qualify as “highly” likely to be true.

In the next two sections the task will be, first, to get clearer on what kind of property coherence is, and, second, to obtain a better understanding of how truth conduciveness should be construed, more precisely. It is implausible to think that coherence is truth conducive in the absence of further conditions: a well-composed novel is usually not true, and yet it may still be highly coherent—perhaps far more so than reality itself. This raises the question of what the additional prerequisites might be, a topic that will be dealt with in sections IV and V. In the final section, I will return to the presystematic question of whether coherence implies truth and consider a different rendering of it.

One of the theses advanced in this paper will be that common criticisms against a connection between coherence and truth are ill-founded, resting on an inadequate and uncharitable understanding of truth conduciveness. But I will also argue that even on a more adequate rendering of that notion, coherence is at best truth conducive in a very weak sense.

## II. THE CONCEPT OF COHERENCE

C. I. Lewis defined coherence—or “congruence”, to use his favored term—as follows (p. 338):<sup>4</sup>

---

<sup>2</sup> See Peter Klein and Ted A. Warfield, “What Price Coherence?”, *Analysis*, LIV, 3 (July 1994): 129-132, and “No Help For the Coherentist,” *Analysis*, LVI, 2 (April 1996): 118-121; Trenton Merricks, “On Behalf of the Coherentist,” *Analysis*, LV, 4 (Oct. 1995): 306-309; Charles B. Cross, “Coherence and Truth Conducive Justification,” *Analysis*, LIX, 3 (July 1999): 186-93; and Tomoji Shogenji, “Is Coherence Truth-Conducive?”, *Analysis*, LIX, 4 (Oct. 1999): 338-45. See also Luc Bovens and Erik J. Olsson, “Coherentism, Reliability and Bayesian Networks,” *Mind*, CIX (Oct. 2000): 685-719, and my “Why Coherence Is Not Truth-Conducive,” *Analysis*, LXI, 3 (July 2001): 236-41.

<sup>3</sup> “What Price Coherence?”, p. 129.

<sup>4</sup> Lewis prefers ‘congruence’ to ‘coherence’ because he wants to mark his departure from British post-Kantian idealism and its “coherence theory of truth.”

A set of statements, or a set of supposed facts asserted, will be said to be congruent if and only if they are so related that the antecedent probability of any one of them will be increased if the remainder of the set can be assumed as given premises.

The set  $S$  consisting of propositions  $A_1, \dots, A_n$  is congruent relative to a probability distribution  $P$  just in case  $P(A_i | B_i) > P(A_i)$  for  $i = 1, \dots, n$ , where  $B_i$  is a conjunction of all elements of  $S$  except  $A_i$ .

There are two reasons why I will choose not to follow Lewis in this respect. The first is trivial: since my concern is with the question whether more coherence implies a higher likelihood of truth, I am interested, first and foremost, in a notion of coherence admitting of (non-trivial) degrees, whereas Lewis is here proposing an absolute (non-graded) conception.<sup>5</sup>

The second reason is more fundamental. As the following example shows, congruence fails even as a reasonable explication of absolute coherence. The reason why it fails is of relevance to our concerns.<sup>6</sup> Suppose that there is a reasonable number of students and a reasonable number of octogenarians (80-89 year olds). Suppose that all and only students like to party and that all and only octogenarians are bird watchers, and that there are some, but very few, octogenarian students. A murder happened in town. Consider the following propositions:

$A_1$  = "The suspect is a student"

$A_2$  = "The suspect likes to party"

$A_3$  = "The suspect is an octogenarian"

$A_4$  = "The suspect likes to watch birds"

The set  $S = \{A_1, A_2, A_3, A_4\}$  is congruent in Lewis's sense but intuitively anything but coherent: one half of the story (the one about the partying student) is very unlikely given the other half (the one about the bird watching octogenarian).

What has gone wrong here? One notable thing is that the joint probability of the propositions in  $S$  is very low; the probability that the suspect is both a partying student and an octogenarian bird watcher is close to zero. This suggests identifying the degree of coherence of a set with its joint probability.

$$C_0(A, B) = P(A \wedge B).$$

This measure takes on a minimum value of 0 if and only if there is no overlap between  $A$  and  $B$ . It takes on a maximum value of 1 just in case  $P(A) = P(B) = 1$ , i.e., just in case both  $A$  and  $B$  are for sure. It is obvious how to generalize this measure to the case of an arbitrary (finite) number of propositions:<sup>7</sup>

$$C_0(A_1, \dots, A_n) = P(A_1 \wedge \dots \wedge A_n)$$

---

<sup>5</sup> Lewis extrapolates a notion of "extent of congruity" (p. 357) from his absolute notion of congruence but leaves the details of this non-trivial undertaking in the dark.

<sup>6</sup> The following counter example is adopted from Bovens and Olsson (*op. cit.*).

<sup>7</sup> Throughout the paper I will be concerned exclusively with the coherence and truth of finite sets.

Suppose, however, that there has been a robbery. To get an unbiased view on who might have committed the crime you decide to consult four different witnesses. Suppose that the first two witnesses both claim that Steve did it, the third witness that Steve, Martin or David did it, and the fourth witness that Steve, John or James did it. Which pair of statements is the more coherent—that delivered by the first two sources or that delivered by the last two sources? It is difficult to escape the feeling that the statements delivered by the first two sources are more coherent in an intuitive sense. After all, unlike the last two witnesses the first two say exactly the same thing, so that the degree of agreement is higher. According to the  $C_0$  measure, by contrast, the degree of coherence is the same, equaling the probability that Steve did it.

This leads us directly to the next proposal. Rather than measuring the overlap of propositions, we might measure the extent to which they agree. The more the propositions agree, the more coherent they are. From this perspective, propositions that coincide are always maximally coherent, even if their joint probability is not very high. A simple way to measure the extent of agreement is the following:

$$C_1(A,B) = \frac{P(A \wedge B)}{P(A \vee B)}$$

$C_1(A,B)$  measures how much of the total probability mass assigned to either  $A$  or  $B$  falls into their intersection.  $C_1(A,B)$  takes on values between 0 and 1. As before, the degree of coherence is 0 if and only if  $P(A \wedge B) = 0$ , i.e., just in case  $A$  and  $B$  do not overlap at all, while the degree of coherence equals 1 if and only if  $P(A \wedge B) = P(A \vee B)$ , i.e., just in case  $A$  and  $B$  coincide. The measure is straightforwardly generalizable:

$$C_1(A_1, \dots, A_n) = \frac{P(A_1 \wedge \dots \wedge A_n)}{P(A_1 \vee \dots \vee A_n)}$$

An alternative measure of coherence was recently introduced by Tomoji Shogenji<sup>8</sup>:

$$C_2(A,B) = \frac{P(A/B)}{P(A)} = \frac{P(A \wedge B)}{P(A) \times P(B)}$$

Like the other measures,  $C_2(A,B)$  equals 0 if and only if  $A$  and  $B$  do not overlap. But while  $C_1$  takes on its maximum value when the propositions coincide regardless of how specific those propositions are, this is not so for  $C_2$ . For suppose that  $A$  and  $B$  coincide. Letting  $x = P(A)$  we get  $C_2(A,B) = P(A \wedge B)/P(A) \times P(B) = x/x^2 = 1/x$ . Hence, the lower the probability of the coinciding propositions is—i.e. the more specific the information is—the higher is the coherence.<sup>9</sup> This measure, too, allows for effortless extension to the general case:

<sup>8</sup> See his “Is Coherence Truth-Conducive?”.

<sup>9</sup>  $C_2$ , of course, is nothing but one of Rudolf Carnap’s confirmation measures discussed at length in Chapter 5 of his *Logical Foundations of Probability* (The University of Chicago Press, 1950). The other Carnapian confirmation measure,  $C(A,B) = P(A|B) - P(A)$ , cannot be used in this context, as it is not symmetric in its arguments.

$$C_2(A_1, \dots, A_n) = \frac{P(A_1 \wedge \dots \wedge A_n)}{P(A_1) \times \dots \times P(A_n)}$$

The difference between  $C_1$  and  $C_2$  stands out most clearly when two pairs of coinciding propositions are contrasted that differ with respect to specificity. Let us modify the robbery example somewhat. Suppose, as before, that the first two witnesses report that Steve did it, and that the third witness reports that Steve, Martin or David did it. But suppose now that the report of the fourth witness coincides with that of the third, so that they both report that Steve, Martin or David did it. Which testimonies are now the most coherent, those of the first two witnesses or those of the last two? Both sets are in full agreement, and so their degree of coherence should presumably be the same. Appealing to our intuitions of mutual support yields the same result; for each set, the one statement in the set is established if the other is assumed as given premise. This is also what we get if we apply  $C_1$ . On the other hand, the agreement between the first two testimonies is surely much more striking since, unlike the last two, they coincide on a very specific statement. Therefore, one could argue, the degree of coherence should also be greater in the first case. This is also what  $C_2$  yields.

There may be many other ways of measuring the degree of coherence or agreement. While both  $C_1$  and  $C_2$  have some initial appeal, I doubt that our intuitions are clear enough to single them out as the only plausible candidates. Nonetheless, they do provide us with a useful starting point for concrete discussion.

### III. PROPOSITIONAL VS. DOXASTIC TRUTH CONDUCTIVENESS

It appears that most philosophers who have discussed coherence and likelihood of truth of sets of propositions have thought of the relevant likelihood as the likelihood that the whole set be true, i.e., the likelihood of *joint* truth.<sup>10</sup> In other words, the relevant probability is the probability of the *conjunction* of all propositions in the relevant set.

This leads me to my first attempt to explicate what it means for coherence to be truth conducive. A measure of coherence will be said to be propositionally truth conducive if and only if a higher degree of coherence implies a higher probability of joint truth.

*Definition 1:* A measure  $C$  of coherence is *propositionally truth conducive* if and only if the following holds: if  $C(A_1, \dots, A_n) > C(B_1, \dots, B_m)$ , then  $P(A_1 \wedge \dots \wedge A_n) > P(B_1 \wedge \dots \wedge B_m)$ .

Note that  $n$  and  $m$  might be distinct; that is, the size of the sets is allowed to vary. The size issue will be addressed at greater length in section V.

---

<sup>10</sup> For two exceptions see Merricks, “On Behalf of the Coherentist”, and, following him, Shogenji, “Is Coherence Truth-Conducive?”, especially p. 344.

Are there any (non-trivial) propositionally truth conducive coherence measures?<sup>11</sup> A measure is truth conducive, in this sense, if more coherence means higher joint probability. According to  $C_0$ , coherence *is* the joint probability, and so this measure, unsurprisingly, comes out as truth conducive in the propositional sense. This fact actually tells against, rather than in favor of, the reasonableness of our explication.  $C_0$  is not a plausible measure of coherence, and it would be highly surprising, therefore, if it turned out to be, in any interesting sense, truth conducive.

More evidence points in the same direction. Both  $C_1$  and  $C_2$ , two initially attractive measures, turn out not to be truth conducive in the sense of definition 1. The (unmodified) robbery example in section II provides us with a counterexample to the propositional truth conduciveness of  $C_1$ . Recall that the first two witnesses claimed Steve to be the culprit, the third that Steve, Martin or David did it and, finally, the fourth that Steve, John or James is to blame. As we noted before, the first two statements are more  $C_1$ -coherent than the latter two, and yet the joint probability of the two sets is the same. Concerning  $C_2$  it was shown in section II that among pairs of coinciding propositions it yields a higher coherence value for more specific pairs. On the other hand, more specific coinciding propositions have a lower (joint) probability.<sup>12</sup>

What conclusions can be drawn from these difficulties? One possible reaction is to conclude that they teach us an important lesson about the connection between coherence and truth, or rather about the lack of such a connection. This is the path taken by Klein and Warfield. Appealing to a detective story and to our intuitions about coherence in that particular case, they claim to have established that coherence is not truth conducive in the above sense.<sup>13</sup> They conclude that coherence theories of knowledge that rely on coherence to be truth conducive are profoundly mistaken, taking the theory of Laurence Bonjour<sup>14</sup> to be a case in point.

Another possible response, which I have already hinted at, is to question the explication of truth conduciveness. Following this track, Shogenji<sup>15</sup> notes, as we have also done, that if coherence is measured using the measure here called  $C_2$ , then a stronger (i.e. more specific) set may well be

---

<sup>11</sup> Measures that do not make discriminations as for coherence are trivially truth conducive since the antecedent of the implication in definition 1 will always be false. Such measures are disregarded in the following.

<sup>12</sup> These counterexamples also disprove the weaker claim that  $C_1$  and  $C_2$  are propositionally truth conducive among sets of propositions of the same size.

<sup>13</sup> See their "What Price Coherence?". The detective story runs as follows (pp. 130-1): A detective has gathered a large body of evidence that provides a good basis for pinning a murder on Mr. Dunnit. In particular, the detective believes that Dunnit had a motive for the murder and that several credible witnesses claim to have seen Dunnit do it. However, because the detective also believes that a credible witness claims that she saw Dunnit two hundred miles away from the crime scene at the time the murder was committed, her belief set is incoherent (or at least somewhat incoherent). Upon further checking, the detective discovers some good evidence that Dunnit has an identical twin whom the witness providing the alibi mistook for Dunnit. Now the extended set is formed by adding the belief about the twin, which is not implied by what was previously believed. It follows by probability calculus that the extended set is less likely to be true than the original, and yet the former is arguably more coherent than the latter. Klein and Warfield conclude that "coherence, per se, is not truth-conducive" (*ibid.*, p. 131).

<sup>14</sup> *The Structure of Empirical Knowledge* (Boston: Harvard University Press, 1985).

<sup>15</sup> See his "Is Coherence Truth-Conducive?".



more coherent although it is less likely jointly to be true, so that  $C_2$ -coherence does not come out as truth conducive (in the above sense). Shogenji now suggests that this conception of truth conduciveness is mistaken. Let us by the *total individual strength* of a set of propositions mean the product of the probabilities of the individual propositions in the set. Thus, the total individual strength of  $\{A_1, \dots, A_n\}$  is  $P(A_1) \times \dots \times P(A_n)$ . The claim is that when assessing the truth conduciveness of coherence “we need to check whether more coherent beliefs are more likely to be true together than less coherent *but individually just as strong beliefs*” (*op. cit.*, p. 342); that is, the only relevant comparisons, on this view, are those among sets having the same total individual strength. What Shogenji proposes, then, is the following: A coherence measure  $C$  is truth conducive if and only if: if  $C(A_1, \dots, A_n) > C(B_1, \dots, B_m)$  and  $P(A_1) \times \dots \times P(A_n) = P(B_1) \times \dots \times P(B_m)$ , then  $P(A_1 \wedge \dots \wedge A_n) > P(B_1 \wedge \dots \wedge B_m)$ . He observes that on this test, which he thinks is “more reasonable” (*ibid.*, p. 342),  $C_2$ -coherence is truth conducive.<sup>16</sup> I have argued elsewhere<sup>17</sup> that Shogenji does not have a good case for fixing the total individual strength, a point which I will return to in section V.

There is also the option of blaming the troubles on our preliminary measures of coherence. But whether or not one ultimately finds these measures unsatisfactory, there is a fundamental reason to be dissatisfied with our explication of truth conduciveness, a reason different from the unconvincing one offered by Shogenji. The crucial observation is that what we are primarily interested in, from an epistemological perspective, is the coherence and truth of *beliefs*, not the coherence and truth of sets of propositions considered in the abstract of a believer.<sup>18</sup> In other words, the interesting question is not whether more coherent propositions are more likely jointly to be true in general, but whether this is so among propositions that are actually held as beliefs.

Presumably, Klein and Warfield would agree that the important question is what holds for believed propositions. For they declare that they want to show “that by increasing the coherence of a set of *beliefs*, the new, more coherent set of *beliefs* is often less likely to be true than the original, less coherent set.”<sup>19</sup> Unfortunately, they lose sight of the distinction between propositions in general and believed propositions in their probabilistic argumentation. By the same token, Shogenji is officially concerned with beliefs, unofficially with bare propositions.<sup>20</sup>

---

<sup>16</sup> Shogenji’s proposal to fix the strength is made in the context of his criticism of Klein and Warfield’s view and does not represent his own approach to the evaluation of truth conduciveness. The position which he finally arrives at is that truth conduciveness should not be evaluated at the level of sets but at the level of individual beliefs. This leads him to conclude that coherence is *not* truth conducive after all because, he argues, “coherence per se has no bearing on the truth of individual beliefs” (*op. cit.*, p. 344).

<sup>17</sup> See my “Why Coherence Is Not Truth-Conducive”.

<sup>18</sup> Hilary Putnam makes a similar observation in his *Pragmatism. An Open Question* (Cambridge: Blackwell, 1995): “[c]oherence theorists have always pointed out that what they require for truth is not mere coherence of *sentences* but coherence of *beliefs*” (p. 64).

<sup>19</sup> “What Price Coherence?”, p. 129, my emphasis.

<sup>20</sup> Shogenji addresses coherence of believed propositions in “The Role of Coherence in Epistemic Justification,” *Australasian Journal of Philosophy*, LXXIX, 1 (March 2001): 90-106.

Let us state the alternative proposal with more precision. We let  $\text{Bel}_S A$  stand for “The subject  $S$  believes that  $A$ ”.

*Definition 2:* A coherence measure  $C$  is *doxastically truth conducive* (for  $S$ ) if and only if: if  $C(A_1, \dots, A_n) > C(B_1, \dots, B_m)$ , then  $P(A_1 \wedge \dots \wedge A_n | \text{Bel}_S A_1, \dots, \text{Bel}_S A_n) > P(B_1 \wedge \dots \wedge B_m | \text{Bel}_S B_1, \dots, \text{Bel}_S B_m)$ .

The definition says that a measure of coherence is doxastically truth conducive just in case a more coherent set of believed propositions is more likely to be true than a less coherent set of believed propositions. The subject  $S$  will usually be omitted. I will refer to  $\text{Bel} A$  as a *belief report* and to  $A$  as the *content* of that report. The difference between the two conceptions of truth conduciveness is that the doxastic conception, unlike its propositional counterpart, is a conditional notion: it conditionalizes on the assumption that the propositions in question are believed (by a given subject). In the next section, I will put forward some conditions that seem required for coherence to be doxastically truth conducive. The remainder of this section is devoted to the elucidation of definition 2.

We could get this far without touching the difficult subject of how to interpret probability statements. However, whether definition 2 makes good sense does depend, to a non-negligible degree, on how probability is construed. Suppose that we decide to adopt a subjective interpretation. On one influential view, fully believing a proposition requires assigning a subjective (personal, credal) probability of 1 to that proposition.<sup>21</sup> Therefore, the probability of  $A_1 \wedge \dots \wedge A_n$ , given that I believe each conjunct, would also be 1. But then the consequent of the defining condition in definition 2 expresses the falsity that  $1 > 1$ , which was obviously unintended and leads to trivial results. Indeed, on a subjective rendering of probability coherence itself is an entirely trivial matter if it is applied to propositions already believed: the coherence of any set of beliefs is 1, regardless of whether coherence is measured by  $C_1$  or  $C_2$ , or, it would seem, any other reasonable measure.

Fortunately, there are independent reasons not to interpret probability subjectively in this context. In order to retain the connection to the original, pre-systematic problem of coherence and (objective) truth, it is customary even among those who advance a propositional conception of truth conduciveness to adopt an objectivist interpretation of probability.<sup>22</sup> I will do so, too. In leaving it open what objective probability means more precisely I also follow what seems to be the mainstream. The conclusion of this paper will eventually be that coherence is at best truth conducive in a weak sense, even if we objectify coherence. It goes without saying that the

<sup>21</sup> For a prominent example, see Isaac Levi, *The Fixation of Belief and Its Undoing* (Cambridge University Press, 1991).

<sup>22</sup> See Klein and Warfield, “What Price Coherence”, p. 130, and Merricks, “On Behalf of the Coherentist”, pp. 308-9.

connection between coherence and truth, if there is such a connection at all, will be no stronger under other, non-objective interpretations. Thus, from the point of view of coherence and truth, we did the coherence theorist a favor when we objectified coherence.<sup>23</sup>

The simple observations which led us to conclude that coherence, in the sense of  $C_1$  or  $C_2$ , is not propositionally truth conducive do not carry over to the doxastic case. When refuting the propositional truth conduciveness of  $C_2$ , for instance, we referred to cases of two pairs of coinciding propositions, one pair more specific than the other. We noted, on the one hand, that  $C_2$  yields a higher coherence value for the more specific case and, on the other hand, that the (joint) probability will in this case be lower. By probability is here meant the *antecedent* joint probability. But from the fact that the antecedent joint probability is lower for the more coherent case, it does not follow that the *posterior* joint probability—the joint probability conditional on belief—is also lower.

Thus, the move from propositional to doxastic truth conduciveness blocks the general inference from a lower antecedent joint probability to a lower likelihood of truth. This notwithstanding, there are cases in which a lower joint probability, and a correspondingly higher degree of  $C_2$ -coherence, is in fact accompanied by a lower likelihood of truth (in the conditional sense). Compare agreement on two tautologies with agreement on two more specific propositions. The  $C_2$ -coherence will be higher in the more specific case, and yet the joint probability in the case of the tautologies will be 1, regardless of whether we are talking about conditional or unconditional probability, and so there is no room for the probability to be higher in the more coherent scenario.<sup>24</sup> The general lesson to be learnt from this observation is that a coherence measure, in order to be doxastically truth conducive, must assign agreement on propositions with antecedent probability 1 maximum degree of coherence.  $C_2$  fails to satisfy this requirement. By contrast,  $C_1$  does satisfy it, taking on its maximum value 1 for any collection of coinciding propositions, tautologies included.

What is particularly noticeable about definition 2 is that it introduces a new propositional layer into the picture. We now have, on the one hand, propositions like  $A_1$ ,  $A_2$ , and so on, with the intended interpretation that they are about the world (or some interesting part of it) and, on the other hand, propositions like  $\text{Bel}A_1$ ,  $\text{Bel}A_2$ , and so on, which are about what someone believes about the world. Coherence pertains to the first level, i.e., to the level of content rather than to the level of belief report. More will be said in the next section about the relation between belief reports and their contents.

---

<sup>23</sup> Strictly speaking, we do not have to adopt an objective interpretation of probability if the only thing we want to do is escape the threat of collapse. So long as *I* am interesting in the coherence and truth of *your* beliefs, I can use my subjective probabilities. However, as just mentioned, the use of subjective probabilities would make a connection between coherence and *objective* truth more problematic than if objective probabilities were employed.

<sup>24</sup> Similar observations are made in Shogenji's "Is Coherence Truth-Conducive?" and my "Why Coherence is not Truth-Conducive."

As was hinted at already in the introduction, although we are primarily interested in the truth conduciveness of coherence as applied to beliefs, the main issues do not pertain essentially to belief reports, but to reports in general.<sup>25</sup> A statement to the fact that a person believes a given proposition can be taken as a report on that proposition, in the sense of (putatively) indicating its truth. For instance, learning that Einstein believed that  $A$ , where  $A$  is some proposition about physics, would make most of us more inclined to think that  $A$  is true. (If in doubt about this particular example, the reader is encouraged to replace Einstein by an authority of his or her own choice.) Belief reports are just one type of reports. Smith's saying that  $A$  or Jones's remembering that  $B$  are examples of other types. In general, we are presented with a number of reports, and we are interested in the relation between the coherence of the report contents and their (conditional) joint probability. The most general question, then, is whether more coherent propositions are more likely jointly to be true among propositions that have been reported individually to be so. It should be mentioned that the sort of interpretative difficulties that arose in connection with belief reports do not arise for reports in general. My subjective probability that  $A$ , given that I believe that  $A$  is true, is always 1, a fact which was seen above to lead to trivialization in combination with the conditional account of truth conduciveness. But my subjective probability that  $A$ , given that I (seem to) remember that  $A$  is true, need not be 1. For similar reasons, subjective vocabulary can be unproblematically adopted in most other contexts involving non-doxastic reports.

In an attempt to spell out the sense of truth conduciveness underlying the epistemology of Laurence Bonjour, Charles B. Cross arrives at a conditional account similar to ours. Following Cross, we let the relation  $J(B, c, s, r)$  stand for " $B$  is justified to a degree defined by  $\langle c, s, r \rangle$ ." The relation holds if and only if " $B$  is the conjunction of the members of the current belief set of an actual agent whose belief history has length  $r$  and consists of belief sets that have remained coherent to degree  $c$  and stable to degree  $s$  while satisfying the Observation Requirement" (Cross, *op. cit.*, p. 189). In order to satisfy Bonjour's Observation Requirement, a system of beliefs "must contain laws attributing a high degree of reliability to a reasonable variety of cognitively spontaneous beliefs" (Bonjour, *op. cit.*, p. 144). The latter are roughly beliefs that have been acquired non-inferentially. Cross goes on to construe Bonjour's truth conduciveness claim as follows: If  $\langle c_2, s_2, r_2 \rangle$  represents a greater degree of justification than  $\langle c_1, s_1, r_1 \rangle$ , then  $P(B_2 | J(B_2, c_2, s_2, r_2)) > P(B_1 | J(B_1, c_1, s_1, r_1))$ .

Note that Cross conditionalizes, just as we did, on the propositions being believed by an actual agent. However, he explicates the truth conduciveness of a certain notion of (degree of) justification seen as applicable to a package of three parameters, including not only the degree of

---

<sup>25</sup> This point is stressed by Lewis (p. 357): "It may also serve to emphasize the importance of congruence in the confirmation of empirical beliefs if we observe in how large a measure the final basis of credibility must be found in evidence having the character of 'reports' of one kind or another—reports of the senses, reports

coherence but also two additional parameters: the degree of stability and the length of the belief history. While Cross may well be right in that this conception corresponds to Bonjour's intentions, he has not answered our question whether, and in what sense, coherence *per se* is truth conducive (nor did he intend to). Moreover, it seems to me sound methodology to study the truth conduciveness of the components before studying the truth conduciveness of a whole package. The Bonjour/Cross strategy thus reverses the natural order of inquiry. What I will say in the next section will raise the suspicion that the Observation Requirement is actually doing all the work in the resulting theory.

#### IV. INDEPENDENCE AND PARTIAL RELIABILITY

It is implausible to think that coherence could be truth conducive, in the doxastic sense, in the absence of further conditions, e.g., conditions connecting the belief reports with their contents. In this section, I will be concerned with the question of what these additional conditions might be. Before we go on to consider some more substantial conditions, we note that the joint probability of the contents of the reports must not be 0 since nothing, coherence being no exception, could affect the joint probability in that case. For the same reason, the joint probability of the contents must not be 1. The satisfaction of these two conditions will be tacitly assumed in the following.

According to C. I. Lewis, a necessary condition for coherence to raise credibility is that the reporters tell their story independently.<sup>26</sup> For if a coherent set is "fabricated out of whole cloth, the way a novelist writes a novel, or if it should be set up as an elaborate hypothesis ad hoc by some theorist whose enthusiasm runs away with his judgment, such congruence would be no evidence of fact" (p. 352). If it were such evidence then "unreliable reporters would be working in the interest of truth if they got together and fudged their stories into agreement" (*ibid.*), a notion which Lewis, quite rightly, dismisses as absurd.

Taking Lewis's remarks on independence as his point of departure, Laurence Bonjour (*op. cit.*, p. 148) writes:

[a]s long as we are confident that the reports of the various witnesses are genuinely independent of each other, a high enough degree of coherence among them will eventually dictate the hypothesis of truth telling as the only available explanation of their agreement... And by the same token, so long as cognitively spontaneous beliefs are genuinely independent of each other, their agreement will eventually generate credibility...

Bonjour is saying that witnesses have to be independent in order for their agreement to have any effect on the probability of what they agree upon. He is adding that the same holds for cognitively spontaneous beliefs. Coherence among such beliefs has a positive effect on their credibility only to

---

of memory, reports of other persons—and this label 'report' is appropriate just because such items do not fully authenticate what is 'reported'."

the extent that they are independent of each other. We recall that cognitively spontaneous beliefs are acquired non-inferentially, which makes it plausible to think that they should indeed normally be independently held, ‘normally’ because we cannot rule out altogether the possibility that there be common causal factors behind such beliefs.

We may safely conclude that coherence is not truth conducive if the reports are entirely dependent on each other; in such cases, more coherence does not imply a higher posterior joint probability.<sup>27</sup> On the other hand, it is implausible to require full independence for coherence to have the desirable effect; intuitively, a tiny influence of the one report on the other does not cancel out the effect of coherence entirely, although it does make that effect less pronounced. Thus, some degree of dependence is compatible with coherence raising the joint probability, but the raise will be less significant than it would have been, had the reports been less dependent.

Let us turn to the reliability issue. One rarely recognized requirement is that the reports must not be fully reliable. That goes for belief reports in particular. If, for each proposition  $A$ , the probability of  $A$ , given that  $S$  believes that  $A$ , is 1, then the joint probability of the contents of  $S$ ’s beliefs will also be 1, however coherent or incoherent those contents are. If fully reliable reports are admitted, then coherence is not truth conducive. Coherence is impotent for fully reliable believers, but they, of course, have no use for it anyway.

Lewis requires the reports to be, as he puts it, “relatively unreliable” (p. 346) for coherence to have any effect on the credibility of the reported information. His choice of term indicates that he wants to rule out, among other things, full reliability. Lewis also intends to disqualify the possibility of a report’s being entirely disconnected from its content. According to Lewis, “[f]or any of the reports taken singly, the extent to which it confirms what is reported may be slight” (p. 346). But it must not be zero; he writes, in his discussion of memory reports, that “[i]f ... there were no initial presumption attaching to the mnemically presented ... then no extent of congruity with other such items would give rise to any eventual credibility” (p. 357).

I prefer ‘partial reliability’ to Lewis’s expression ‘relative unreliability’ and will use the former in the following. Thus a partially reliable report lies strictly between full reliability and irrelevance. Coherence, according to this view, is effective only if, for each report  $E$  with content  $A$ ,  $1 > P(A|E) > P(A)$ .<sup>28</sup> Applying this to belief reports, we should require that  $1 > P(A|BelA) > P(A)$ ; that is, we should require that the presence of a belief support the content of the belief, though without fully authenticating that content.

---

<sup>26</sup> See, for example, p. 346.

<sup>27</sup> The case of witnesses copying other witnesses’ statements is discussed in Alvin Goldman, “Experts: Which Ones Should You Trust?”, *Philosophy and Phenomenological Research*, LXIII, 1 (July 2001): 85-110. Goldman shows, using Bayesian reasoning, that the addition of such witnesses has no effect on the probability of what is being said.

<sup>28</sup> Note that  $1 > P(A|E) > P(A)$  is equivalent to  $P(E|A) > P(E|\neg A) > 0$ , provided that  $1 > P(A) > 0$ .

We have seen that Lewis, once again, is giving expression to the common sense view: coherence of completely useless information does not do a thing for the credibility of that information. Coherence cannot create credibility out of nothing, but can at best amplify credibility that is already there. This notwithstanding, Lewis's position on reliability has been contested, indeed by Bonjour, who writes: "[w]hat Lewis does not see, however, is that his own [witness] example shows quite convincingly that no antecedent degree of warrant or credibility is required" (*op. cit.*, p. 148). On one plausible interpretation, Bonjour is saying that coherence is truth conducive also when the reports, taken individually, are irrelevant to what is reported.<sup>29</sup>

In an attempt to resolve the dispute between Lewis and Bonjour, I will give a formal argument to the effect that, unlike what Bonjour claims, independence and complete unreliability are not sufficient for coherence to be truth conducive.<sup>30</sup> More specifically, I will argue that full independence should be given a certain exact interpretation and then show that coherence fails to be truth conducive under the assumptions of full independence and complete unreliability. I have argued elsewhere that independence, in this context, should be explicated as conditional independence.<sup>31</sup> What follows is a detailed defense of this proposal.

The idea is that reports are independent just in case they are directly influenced only by the respective facts they report on: they are not directly influenced by what other reporters have to say, nor are they directly influenced by other facts than the facts they report on. It will be helpful to refer to an example throughout this section. (It will be clear, I hope, that nothing hinges on the peculiarities of the example.) Suppose there has been a robbery with Robert's being is one of the suspects. There are two witnesses, Helen and Peter, available for questioning. Consider the following propositions:

$A_1$  = Robert has a gun

$A_2$  = Robert is a professional criminal

$E_1$  = Helen says that Robert has a gun

$E_2$  = Peter says that Robert is a professional criminal

Applying the foregoing explanation,  $E_1$  and  $E_2$  are independent reports on  $A_1$  and  $A_2$ , respectively, just in case  $E_1$  is directly influenced only by  $A_1$ , and  $E_2$  is directly influenced only by  $A_2$ . This does

---

<sup>29</sup> Alternative interpretations of Bonjour's claim are discussed in Bovens and Olsson (*op. cit.*).

<sup>30</sup> What follows is a generalization of an argument due to Michael Huemer. See his "Probability and Coherence Justification," *The Southern Journal of Philosophy*, XXXV, 1997: 463-472. Huemer proves a similar result for the simpler setting in which both witnesses say the same thing.

<sup>31</sup> Bovens and Olsson (*op. cit.*). For more on conditional independence, see Wolfgang Spohn, "Stochastic Independence, Causal Independence, and Shieldability," *Journal of Philosophical Logic*, IX, 1980: 73-99. Underlying the theory of Bayesian networks, a currently expansive branch of computer science, conditional independence is already a well-established concept of independence. For a standard text on Bayesian

not mean that there could not be other, indirect influences among the propositions. Conditional independence is compatible with  $E_1$ 's being influenced also by other propositions than  $A_1$ , in which case, however, the influence must pass via  $A_1$  and hence be indirect.

How do we translate this into probability theory? The trick is to block the communication path to  $E_1$  via  $A_1$  and see if there are still any influences on  $E_1$ . There are two ways to block that path: assuming  $A_1$  true or assuming  $A_1$  false (or both), i.e. conditioning on  $A_1$  or conditioning on  $\neg A_1$  (or both). Similarly for  $E_2$  and  $A_2$ . Conditional independence is satisfied if there are no residual influences. In the interest of avoiding technicalities, I will confine myself in the following to the two report case.

*Definition 3:*  $E_1$  and  $E_2$  are *independent reports* on  $A_1$  and  $A_2$  if and only if the following hold:

$$\begin{array}{ll}
P(E_2|A_1, A_2) = P(E_2|A_1, A_2, E_1) & P(E_2|\neg A_1, A_2) = P(E_2|\neg A_1, A_2, E_1) \\
P(E_2|A_1, A_2) = P(E_2|A_2) & P(E_2|\neg A_1, A_2) = P(E_2|A_2) \\
P(E_1|A_1, A_2) = P(E_1|A_1) & P(E_1|\neg A_1, A_2) = P(E_1|\neg A_1) \\
\\ 
P(E_2|A_1, \neg A_2) = P(E_2|A_1, \neg A_2, E_1) & P(E_2|\neg A_1, \neg A_2) = P(E_2|\neg A_1, \neg A_2, E_1) \\
P(E_2|A_1, \neg A_2) = P(E_2|\neg A_2) & P(E_2|\neg A_1, \neg A_2) = P(E_2|\neg A_2) \\
P(E_1|A_1, \neg A_2) = P(E_1|A_1) & P(E_1|\neg A_1, \neg A_2) = P(E_1|\neg A_1)
\end{array}$$

Note that independence thus defined is a four-place relation.

According to the second equation on the left, for instance, if we already know that Robert is a criminal, then hearing in addition that he owns a gun does not change our confidence that Peter will report that Robert is a criminal. Our level of confidence that Peter will deliver such a report is determined once the truth value of its content is known; all other factors have been screened off. As a special case, substituting  $\text{Bel}A_1$  and  $\text{Bel}A_2$  for  $E_1$  and  $E_2$  respectively throughout definition 3 gives us an account of when two belief reports are independent with respect to their contents. On this reading, the same equation says that if we already know that Robert is a criminal, then learning in addition that he owns a gun does not change our confidence that Peter will *believe* that Robert is a criminal. Our level of confidence that Peter will adopt such a belief is determined once the truth value of the content is known.

There are two other concepts of independence that should be sharply distinguished from conditional independence in the sense just defined. First, one could by independence mean that the *contents* of the reports are independent, in the sense of  $P(A_1|A_2) = P(A_1)$ . In the Robert case, the

---

networks, see Judea Pearl, *Probabilistic Reasoning in Intelligent Systems* (San Francisco: Morgan Kaufmann, 1988).



contents would be independent, in this sense, if knowing Robert to be a professional criminal would neither add to, nor subtract from, our expectation that he will own a gun. Independence in this sense interferes with coherence and results generally in a low coherence value. For instance, while the  $C_2$  measure can take on any real value greater than or equal to 0, the  $C_2$  value of two content independent propositions is always 1. If content independence were the relevant notion of independence for the purposes of coherence and truth, it would be impossible for two independent witnesses to give agreeing testimonies, which is absurd.

Second, there is independence in the sense of  $P(E_1|E_2) = P(E_1)$ : knowing that Peter has testified that Robert is a criminal does not provide any information whatsoever as to whether Helen will testify that Robert has a gun. Normally, however, hearing one witness testifying to a certain effect raises our confidence that the next witness to be questioned will give an agreeing statement. This kind of dependence is quite compatible with the witnesses being independent in the sense relevant for the purposes of coherence and truth, or so I claim. To see this, suppose that  $E_1$  and  $E_2$  are partially reliable reports on  $A_1$  and  $A_2$ , respectively. Even if Helen and Peter base their reports directly only on the respective facts and have never communicated, Peter's testifying that Robert is a professional criminal would raise our expectations that Helen will testify that Robert has a gun. Why? Peter's being partially reliable, his testimony would raise our confidence that Robert actually is a criminal which, in turn, would raise our confidence that Robert actually has a gun (because of the coherence of criminality and gun possession). Helen's being also partially reliable, it now becomes more probable than it was before that she will testify to the effect that Robert has a gun. Hence, we would be slightly more inclined to expect that Helen will testify that Robert has a gun after we have heard Peter's positive testimony than we were before. And yet the testimonies were assumed to be independent in the relevant sense.

I hope to have made it plausible that the relevant sense of (full) independence is conditional independence and proceed now to the second part of my formal argument for requiring partial reliability, which amounts to showing that coherence is not truth conducive under conditional independence and complete unreliability. Fortunately, this is quickly accomplished.

*Observation:* Suppose that the following hold:

- (i)  $E_1$  and  $E_2$  are independent reports on  $A_1$  and  $A_2$ .
- (ii)  $P(A_1 | E_1) = P(A_1)$  and  $P(A_2 | E_2) = P(A_2)$ .
- (iii)  $A_1 \wedge A_2$ ,  $A_1 \wedge \neg A_2$ ,  $\neg A_1 \wedge A_2$  and  $\neg A_1 \wedge \neg A_2$  all have non-zero probability.

Then  $P(A_1 \wedge A_2 | E_1, E_2) = P(A_1 \wedge A_2)$ .<sup>32</sup>

---

<sup>32</sup> Proof. By Bayes' Theorem,

$$(1) \quad P(A_1 \wedge A_2 | E_1, E_2) = \frac{P(E_1 \wedge E_2 | A_1, A_2) P(A_1 \wedge A_2)}{P(E_1 \wedge E_2)}$$

The third condition serves to rule out certain uninteresting limiting cases. Thus, it follows (essentially) already from conditional independence that reports that are completely unreliable regarding their contents, when taken singly, fail to have any effect on the joint probability of those contents, when combined.

Now, if  $\text{Bel}A_i$  is substituted for  $E_i$  in the observation, it follows that doxastic truth conduciveness collapses into propositional truth conduciveness under (essentially) the conditions of conditional independence and complete unreliability. Having established, in section III, that coherence fails to be propositionally truth conducive, we may conclude that coherence fails to be doxastically truth conducive under (essentially) those conditions. Hence, besides independence, partial reliability is also needed to provide the context within which coherence can do useful, amplifying work.

C. I. Lewis suggested an explication of independence strikingly similar to, but not identical with, the one proposed here. Lewis introduces his concept of independence when discussing the corroboration of a hypothesis through its testable consequences. Since the testable consequences of a hypothesis are (putative) indicators of its truth and hence are reports, in the abstract sense, on that hypothesis, his remarks on independence are easily generalized beyond the hypothesis-consequences setting.<sup>33</sup> He considers two propositions,  $A$  and  $B$ , to be independent consequences of a hypothesis  $H$  just in case  $A$  and  $B$  are “so related that supposing  $H$  false, the finding of one of them true would not increase the probability of the other” (p. 344). Lewis restates this idea several times, with little variation, writing for instance that “in general, the consequences of a hypothesis are independent only in the sense that the establishment of one does not increase the probability of another on the assumption that the hypothesis is *false*” (p. 349, footnote 6, original emphasis). It was hence clear to Lewis that the relevant notion of independence is of a conditional nature. However, Lewis insisted on taking into account only independence statements conditional on the

---

By conditional independence,  $P(E_1 \wedge E_2 | A_1, A_2) = P(E_1 | A_1)P(E_2 | A_2)$ . It follows from (ii) and familiar probabilistic facts that  $P(E_i | A_i) = P(E_i)$ ,  $i = 1, 2$ . Hence,

$$(2) P(E_1 \wedge E_2 | A_1, A_2) = P(E_1)P(E_2)$$

By (iii) and the theorem of total probability,  $P(E_1 \wedge E_2) = P(E_1 \wedge E_2 | A_1, A_2)P(A_1 \wedge A_2) + P(E_1 \wedge E_2 | A_1, \neg A_2)P(A_1 \wedge \neg A_2) + P(E_1 \wedge E_2 | \neg A_1, A_2)P(\neg A_1 \wedge A_2) + P(E_1 \wedge E_2 | \neg A_1, \neg A_2)P(\neg A_1 \wedge \neg A_2)$ . By conditional independence, the right hand side of that equation equals  $P(E_1 | A_1)P(E_2 | A_2)P(A_1 \wedge A_2) + P(E_1 | A_1)P(E_2 | \neg A_2)P(A_1 \wedge \neg A_2) + P(E_1 | \neg A_1)P(E_2 | A_2)P(\neg A_1 \wedge A_2) + P(E_1 | \neg A_1)P(E_2 | \neg A_2)P(\neg A_1 \wedge \neg A_2)$ . As already noticed, it follows from (ii) that  $P(E_1 | A_1) = P(E_1)$  and  $P(E_2 | A_2) = P(E_2)$ . It also follows from (ii) that  $P(E_2 | \neg A_2) = P(E_2)$  and  $P(E_1 | \neg A_1) = P(E_1)$ . Combining all this yields

$$(3) P(E_1 \wedge E_2) = P(E_1)P(E_2)P(A_1 \wedge A_2) + P(E_1)P(E_2)P(A_1 \wedge \neg A_2) + P(E_1)P(E_2)P(\neg A_1 \wedge A_2) + P(E_1)P(E_2)P(\neg A_1 \wedge \neg A_2) = P(E_1)P(E_2).$$

It follows from (1), (2) and (3) that  $P(A_1 \wedge A_2 | E_1, E_2) = P(A_1 \wedge A_2)$ , which ends the proof.

<sup>33</sup> Lewis explains the relevant notion of consequence as follows (p. 343, notation adapted): “when a hypothesis  $H$  is said to have consequences  $C$ , what typically is meant is that  $H$ , together with other statements which may reasonably be assumed, gives a high probability of  $C$ .” He adds, in footnote 5 on the same page, that “[i]t is not even essential that such a probability be high”, that is to say, “in general the same principle will apply wherever  $C$  is more probable than not.”

falsity of the hypothesis, leaving out independence statements conditional on its truth. While Lewis was on the right track, the kind of independence he arrived at is ultimately too weak for the purposes of coherence and truth.

## V. TRUTH CONDUCTIVENESS CETERIS PARIBUS

In this section, I will discuss yet another condition that must be satisfied for coherence to stand a reasonable chance of being truth conducive. Suppose that we want to evaluate the truth conduciveness of a given measure of coherence. We now find that the propositions reported by one set of witnesses are more coherent, according to the measure, than the propositions reported by another set of witnesses, and also that the conditions we have discussed so far are satisfied; that is, in each scenario the witnesses are partially reliable and independent. However, because the witnesses delivering the more coherent set are less reliable, or more dependent, than the witnesses delivering the less coherent set, the posterior joint probability turns out to be higher for the latter. Should we take this as a counterexample to the truth conduciveness of the coherence measure in question? Presumably, we should not. Supposedly, the claim that more coherence implies a higher likelihood of truth is a *ceteris paribus* claim: more coherence implies a higher likelihood of truth *if everything else is equal*. In particular, neither the degree of reliability nor the degree of independence should be allowed to vary across scenarios under comparison.

Shogenji, too, notes that certain “confounding factors” have to be kept fixed in the evaluation of the truth conduciveness of coherence.<sup>34</sup> He does not mention reliability and independence in this connection, which is not surprising considering the fact that he, as we have seen, is concerned with the propositional conception of truth conduciveness and not with the doxastic or testimonial conception. In the propositional one-layer model, issues of reliability and independence (in the relevant sense) do not arise. Instead, Shogenji claims that total individual strength is a confounding factor with respect to coherence. In the evaluation of the truth conduciveness of the latter, the former has to be held fixed.

Before evaluating Shogenji’s contention about strength, let us take a look at *ceteris paribus* conditions from an abstract point of view. As Robert L. Frazier<sup>35</sup> notices, other things are equal when there is only one candidate for being an influence on some outcome. By requiring other influences to be excluded, we are trying to isolate the target property. The truth of a statement with a *ceteris paribus* clause depends on what happens when this isolation is achieved. Frazier (*ibid.*, p. 119) observes that the properties that have been thus singled out must be capable of independent variation. Otherwise they are not really separable. He goes on to describe what it would mean, more precisely, for two properties  $P_1$  and  $P_2$  to be *incapable* of independent variation. That would,

---

<sup>34</sup> See his “Is Coherence Truth-Conducive?”.

<sup>35</sup> See p. 14 in his “Moral Relevance and *Ceteris Paribus* Principles,” *Ratio*, VII, 2 (Sept. 1995): 113-127.

he writes, mean that whenever an object has  $P_1$  to a particular degree, say  $n$ , the object also has  $P_2$  to a particular degree, say  $m$ , so that “when one occurs to a particular degree this is then a certain sign that the other occurs to a particular degree” (*ibid.*).

While I agree with Frazier that the requirement of independent variation would be violated in this case, I do not think that this sort of situation represents the only type in which it would be violated. Rather, this would be an extreme case. For suppose that whenever an object has  $P_1$  to a particular degree, say  $n$ , it has  $P_2$  either to a degree  $m$  or to a degree  $m'$ . This would, intuitively, also amount to lack of independent variation, provided, of course, that  $P_2$  could take on more than just the two values  $m$  and  $m'$  before the value of  $P_1$  was fixed. In general the requirement of independent variation is violated, I take it, if fixing the value of the one property imposes limitations on the extent to which the other property can consistently vary.<sup>36</sup>

In the present case, we are trying to isolate one property, namely coherence, from other properties that may have an influence on the joint probability of a set of beliefs or testimonies. It can now be seen that if we use Shogenji’s own measure of coherence ( $C_2$ ), we cannot include strength among the properties that should be held fixed. The reason is that strength and coherence are not independent properties in the relevant sense. The reader might have suspected this already on the grounds that the strength figures as the denominator in the definition of that measure. To prove the case, note that if the strength, i.e.  $P(A) \times P(B)$ , is not fixed,  $C_2(A, B)$  has no upper bound; it can take on any real value greater than, or equal to, 0. However, if we keep the strength fixed,  $C_2(A, B)$  becomes bounded. For instance, if the strength is assigned a value of 9/10, then the upper bound is 10/9. Thus, keeping the strength fixed imposes limitations on the extent to which the degree of coherence can consistently vary, in violation of the requirement of independent variation.<sup>37</sup>

---

<sup>36</sup> Independent variation, in this sense, is compatible with the instantiation of one property to a particular degree making the instantiation of another property to a particular degree highly likely, so long as the resulting likelihood is less than 1.

<sup>37</sup> In Bovens and Olsson (*op. cit.*, p. 691) we argued that the antecedent joint probability should be held fixed in assessing whether coherence is truth conducive, referring to the following example. Let us assume that two independent witnesses testify to the effect that a dog died due to witchcraft, what was antecedently very improbable. Suppose also that two other independent witnesses provide us with reports that are not in complete agreement but still not mutually exclusive, e.g., that the dog died due to cancer and due to pneumonia, respectively. All reporters are initially considered partially reliable to the same degree. Clearly the former reports are more coherent than the latter reports. And yet, the posterior (joint) probability that the dog died due to witchcraft is still lower than that it died due to the combined effect of cancer and pneumonia. Why? Because of the difference in priors. This example was taken to illustrate the thesis that the prior joint probability may influence the posterior joint probability and that we need to keep it fixed on the grounds that we need to keep fixed all factors, other than coherence, that may affect the posterior joint probability. However, the question is whether the prior joint probability is clearly separable from coherence, so that we may legitimately speak of them as two distinct factors. For reasons similar to those advanced here against fixing the strength, I do not think any more that prior joint probability can be isolated from coherence. Fixing the prior may lead to violations of the requirement of independent variation. The effect of varying the prior probability of what is being said on its posterior probability is studied in my “Corroborating Testimony, Probability and Surprise” (to appear in *The British Journal for the Philosophy of Science*) in the simpler context involving two witnesses testifying to the same effect.

By contrast, holding the degree of reliability or independence fixed is in perfect compliance with the requirement of independent variation. Fixing these aspects of scenarios does not mean limiting in any way what degree of coherence can be consistently attributed. For instance, if we learn that the each witness tells the truth 60 percent of the time, this information does not put us in a better position than we were before to eliminate coherence values.<sup>38</sup>

It is not surprising that coherence and reliability should admit of independent variation: while the former is a property on the content level, the latter concerns the relation between contents and reports. Hence, the former is, in a sense, “horizontal” and the latter “vertical.” Coherence and total individual strength, on the other hand are both properties on the horizontal content level, which is why we should expect them to interfere, as indeed they do.

Should we hold the size of an information set fixed in the evaluation of the truth conduciveness of coherence? That would be reasonable if size were something that could be isolated from coherence. To be sure, just by focusing on information sets of a certain size, one has not thereby in any way restricted the assignment of a degree of coherence. On the other hand, when the size is changed as the result of the addition of one or more items of information (e.g., further testimonies), this usually also affects the degree of coherence. Size is, in this sense, not completely separable from coherence. By contrast, changing the reliability or independence of the witnesses never brings with it a change in coherence. It seems to me that the *ceteris paribus* conception requires this strong sense of separation illustrated by reliability and independence. If this is true, size should not be held fixed.

If size should not be held fixed, then  $C_1$  is not truth conducive *ceteris paribus*. Suppose that we have bet money on the outcome of casting a die but are unable to be present to confirm the result. We decide to consult informants who were present. Consider two scenarios: In the first scenario, two persons were present to watch the event, each reporting that ‘One’ came up. In the second scenario, 100 persons were present, out of which 99 report that ‘One’ came up, and the one remaining that either ‘One’ or ‘Two’ did. In the first scenario, there is perfect agreement among the witness statements (and hence a maximum  $C_1$ -value), in the second not (and hence a less than maximum  $C_1$ -value). Yet it is clear that we would, everything else equal, be more confident that ‘One’ came up in the second, less coherent scenario. Hence, a lower degree of  $C_1$ -coherence may be counterbalanced by a larger number of testimonies. This casts doubts on the doxastic case as well, as it raises the suspicion that a higher degree of  $C_1$ -coherence could be outweighed by a sufficient increase in the number of beliefs.

---

<sup>38</sup> If we were to learn that the witnesses *always* tell the truth, that would indeed tell us something conclusive about the coherence: it would enable us to exclude the case of their delivering incompatible testimonies. Hence, it would seem that reliability is not independent of coherence after all. Recall, however, that in requiring the witnesses to be partially reliable, we have already ruled out perfectly reliable witnesses from consideration.

This negative result depends on the particular coherence measure. If instead  $C_2$  is used, for example, we cannot conclude that the degree of coherence (agreement) is greater in the first scenario (with two reports) than in the second (with 100 reports). In fact, the second scenario will exhibit a much higher degree of  $C_2$ -coherence than the first. To see this, recall the quotient in the definition of  $C_2$ . While the numerator of that quotient (i.e. the joint probability) will be the same for the two scenarios, the denominator (i.e. the total individual strength) will be much smaller for the second scenario, thus making the value of the whole expression much higher.

The problem of coherence and truth, or at least one clear and manageable sub-problem, is finding a (non-trivial) measure of coherence that is doxastically truth conducive *ceteris paribus* under (essentially) the conditions of partial reliability and independence. At the same time, it must be added that there are reasons for pessimism concerning the existence of such measures. The measures we started out with seemed to capture important intuitions about coherence reasonable well but they turned out, nonetheless, not to be truth conducive in this sense. Neither  $C_1$  nor  $C_2$  is truth conducive *ceteris paribus* under the conditions in question.  $C_1$ , as we just observed, has a problem of size. We have noticed before, in section IV, that  $C_2$  has difficulties with strength (specificity).

## VI. ON THE POSSIBILITY OF READING OFF THE POSTERIOR FROM THE COHERENCE

The kind of relation between coherence and truth that has been the subject of our discussion so far is of a comparative nature. Given that a set  $S$  is more coherent than a set  $S'$ , can we conclude that  $S$  is also more likely than  $S'$  to be true as a whole? We found that this question could hardly be answered in the affirmative unless it is qualified in certain substantial respects. In this section, I will try to make likely that we could not hope for a more ambitious connection between coherence and truth than this. In particular, it is implausible to think that there is something like a degree of coherence that would allow us to read off the likelihood of truth of a given set from its degree of coherence. While it remains an open question whether a *higher* degree of coherence implies a *higher* likelihood of truth (given the appropriate qualifications), we can be confident that a *high* degree of coherence does not necessarily imply a *high* likelihood of truth. Once again, my point of departure will be the view of C. I. Lewis.

In his attempt to provide a validation of memory, Lewis argues that we can know, on a priori grounds, that the partial reliability requirement is satisfied in that context.<sup>39</sup> But this is also all we

---

<sup>39</sup> According to Lewis, mnemonic representation affords some probability of past fact already before any examination as to coherence. Such credibility “is constitutive of our sense of the only reality by reference to which empirical judgments could have either truth or falsity or any meaning at all” (p. 361). Lewis’s point is that without this assumption our very notion of empirical fact would not make sense, and so there could be no (empirical) basis from which this assumption could be questioned. The assumption in question “is an analytic statement which can only be repudiated on premises which already imply it” (*ibid.*), and skepticism about its validity is ultimately self-defeating.

can say. For all we know, the initial credibility (reliability) of our memory may be high, low or anything in between. Lewis expresses this contention by saying that “the degree of this initial credibility ... is hardly assignable” (p. 357). According to Lewis, then, our initial condition as regards the reliability of our memory is one of partial ignorance.

Lewis goes on to advance two theses about the relation between coherence and likelihood of truth. Unfortunately, what says in this connection is perplexing, as he appears to contradict himself.

Lewis claims that it does not matter that the degree of initial credibility is not assignable because—and this one of his two theses—“it does not need to be assigned” (p. 357). Now, whether or not something needs or does not need to be done is relative to some purpose. This raises the question of what purpose Lewis could have had in mind here. The answer seems clear enough: the initial credibility does not need to be assigned for the purpose of determining the posterior (eventual) probability of a given set of reported propositions. It is sufficient, for that purpose, to know the degree of coherence pertaining to that set. Thus, the claim is that we can read off the posterior from the degree of coherence, without having to take the initial credibility of the information into account, so long as we know that this credibility is positive.

This thesis is also stated as follows (p. 352, my emphasis):

Take the case of the unreliable observers who agree in what they report. In spite of the antecedent improbability of any item of such report, when taken separately, it may become practically certain, in a favorable case, *merely* through congruent relations to other such items, which would be similarly improbable when separately considered.

Here is another passage in the same spirit (p. 354, my emphasis):

*All that is needed* is initial assumption that the mere fact of present rememberings renders what is thus memorially present in some degree credible. For the rest, the congruence of such items with one another and with present sense experience will be capable of establishing an eventual high credibility, often approximating to certainty, for those items which stand together in extensive relations of such congruence.

Lewis is here clearly committed to the thesis that given the initial credibility of our memories, a high enough degree of coherence (“congruence”) will *by itself* establish a high posterior probability (“eventual high credibility”).

And yet Lewis also writes, in effect, that the posterior probability does *not* only depend on the degree of coherence (p. 357, my emphasis):

A larger or a smaller such initial probability would have no appreciable effect upon the eventually determinable probabilities in question *beyond that of a difference in the extent of congruity with other mnemonic items and with sense presentation which would be required for building up eventual probabilities sufficient for rational and practical reliance.*

What Lewis is here mentioning in passing, as if it were of no particular significance to the issue at hand, is actually the crucial thing. For he is saying that the posterior probability is a function not only of the degree of coherence (“extent of congruity”) but also of the level of initial credibility. More precisely, the lower the initial credibility is, the higher is the degree of coherence required for building up posterior probabilities sufficient for “rational and practical reliance.” If our memories are initially already rather reliable, they do not have to agree very much in order for us to be willing to rely on them. If, on the other hand, they are initially highly dubious, striking agreement would be needed to achieve the same effect. All this presupposes, of course, that the independence condition is satisfied.

Lewis’s second thesis is clearly the correct one. In fact, we took its validity for granted in our discussion of *ceteris paribus* clauses when we argued, in section V, that the degree of reliability should be held fixed in the evaluation of the truth conduciveness of coherence. This has grave consequences for Lewis’s epistemological enterprise. Since the posterior does not depend only on the degree of coherence but also on the initial credibility, in order to know whether we can, rationally and practically, rely on the information at hand, we have to know what the initial credibility is more precisely; it does not suffice just to know that it is positive. Otherwise, we could not be confident that the degree of coherence pertaining to the given information set is high enough. Hence, for the purposes of determining the posterior probability, the initial credibility *does* need to be assigned. But, says Lewis, it cannot be assigned. It follows that, unless the antecedent joint probability was already 1, we can never be confident that our probabilities are sufficient for rational and practical reliance. The net result is a form of skepticism that would have been unacceptable to Lewis.<sup>40</sup>

Apart from their ramifications for Lewis’s theory, these considerations strongly suggest, if they do not prove, that there is no measure, probabilistic or not, definable solely in terms of the contents of beliefs, memories or testimonies that would allow us to conclude anything interesting about how high the posterior probability is in absolute, non-comparative terms. In order to draw such conclusions we also need to know (at least) the initial reliability of the reports. In the absence of such information, we may at best draw the comparative conclusion that more coherence implies a higher posterior probability.

We have already seen, in section IV, that Bonjour’s coherence theory is based on the false probabilistic assumption that individually useless data can, if they are independent, become collectively useful, in the sense of yielding, when combined, a high posterior probability. Lewis wisely rejected this notion, pointing out that coherence has an effect only if the individual data

---

<sup>40</sup> Lewis’s attempted justification of memory beliefs is studied in greater detail in a forthcoming paper by me and Tomoji Shogenji entitled “C. I. Lewis on Congruence and Probability.”



have some degree of initial credibility. But Lewis, too, was unrealistically optimistic in his assessment of what coherence can actually accomplish. His (weak) foundationalism—his attempt to provide a “final validation of empirical knowledge by reference to ultimate data in some sense presently given, and to the congruence of such data” (p. 356), given our supposed partial ignorance regarding the initial credibility of those data—shares the deficiency of Bonjour’s theory of being founded on an incorrect probabilistic supposition. In Lewis’s case, the erroneous, but essential, assumption is that posterior probability is determined by coherence alone once the minimum conditions of partial reliability and independence are satisfied.<sup>41</sup>

Summing up, we have seen that theories that rely on the truth conduciveness of coherence are frequently misinterpreted as advancing a propositional conception of the truth connection. Instead, truth conduciveness should be understood in a doxastic, or testimonial, sense. The general conclusion after this initial confusion has been sorted out is that there is nevertheless at best a very weak connection between coherence and (likelihood of) truth. In particular, we should not hope to establish anything more substantial than the comparative statement that more coherence implies a higher likelihood of truth. Moreover, the comparative statement, in order to stand a reasonable chance of being valid, should be further qualified as involving a *ceteris paribus* clause and as presupposing the obtaining of certain favorable circumstances, mainly partial reliability and independence. Whether we can establish even this much remains an open question. We also saw that Laurence Bonjour’s coherence theory as well as C. I. Lewis’s weak foundationalism turn out to be untenable already on the grounds that they rely, each in its own essential way, on probabilistic assumptions that are simply incorrect.

ERIK J. OLSSON

University of Constance

---

<sup>41</sup> According to Huemer (*op. cit.*, p. 468), “Lewis’ fairly modest claim seems to have been just this: that you can get a high degree of confirmation as a result of the agreement of the witnesses even when each witness has only a low degree of credibility considered by himself.” Huemer takes the truth of the “modest claim” to warrant the conclusion that “weak foundationalism should be taken seriously as an epistemological position” (p. 471). Lewis, indeed, claimed that it is possible to obtain a high degree of confirmation in this way, if the circumstances are favorable. However, unlike what Huemer seems to think, this fairly uncontroversial statement, as I hope to have made clear above, does not exhaust Lewis’s weak foundationalism.