LUC BOVENS and ERIK J. OLSSON

# BELIEVING MORE, RISKING LESS: ON COHERENCE, TRUTH AND NON-TRIVIAL EXTENSIONS

## 1. INTRODUCTION

If you believe more things you thereby run a greater risk of being in error than if you believe fewer things. From the point of view of avoiding error, it is best not to believe anything at all, or to have very uncommitted beliefs. But considering the fact that we all in fact do entertain many specific beliefs, this recommendation is obviously in flagrant dissonance with our actual epistemic practice. Let us call the problem raised by this apparent conflict the Addition Problem.

There are, conceivably, different ways to come to grips with the Addition Problem. One could, for instance, argue that shunning error is not the only epistemic goal; there is also the goal of seeking truth, i.e., true information. Following William James (1897), it is often pointed out that we should see "shun error" and "seek truth" as two competing desiderata.[1] Since the goals are competitors, it is necessary to make a trade-off between them. The result will normally be a belief system situated somewhere strictly between "believing nothing" and "believing everything", its exact position depending on the inquirer's degree of caution.

In this paper we will find reasons to reject another premise used in the formulation of the Addition Problem, namely, the fundamental premise according to which believing more things increases the risk of error. As we will see, acquiring more beliefs need not decrease the probability of the whole, and hence need not increase the risk of error. In fact, more beliefs can mean an *in*crease in the probability of the whole and a corresponding *de*crease in the risk of error.

We will consider the Addition Problem as it arises in the context of the coherence theory of epistemic justification, while keeping firmly in mind that the point we wish to make is of epistemological importance also outside the specific coherentist dispute. The problem of determining exactly how the probability of the whole system depends on such factors

*Erkenntnis* **57:** 137–150, 2002.

as coherence, reliability and independence will be seen to open up an interesting area of research in which the theory of conditional independence structures is a helpful tool.[2]

## 2. COHERENTISM AND THE ADDITION PROBLEM

Advocates of the coherence theory of epistemic justification want coherence to be truth conducive: the more coherent a belief system is, i.e., the more its constituent beliefs hang together, the more likely it is that the beliefs are all true. If coherence were not truth conducive in this sense, it would be difficult to maintain that it has any connection with epistemic justification.[3] In their 1994 article, Klein and Warfield (K&W) present an argument to the effect that coherence lacks the desirable property. In a second paper (K&W 1996) they defend their argument against the criticism of Merricks (1995). We will argue that one of the crucial premises in K&W's argument is false. The underlying error is that K&W do not properly distinguish between belief systems and mere sets of propositions.

Let us consider K&W's argument. The coherentist is committed to the truth conduciveness of coherence:

(1) The more coherent a belief system is, the more probable it is.

Let an extension B′ of a belief system B be non-trivial if some of the beliefs that are in B′ but not in B neither follow logically from B nor have a probability of 1. K&W's argument against (1) rests on the following premises:

(2) Any non-trivial extension of a belief system is less probable than the original system.

(3) There exist non-trivial extensions of belief systems that are more coherent than the original belief system.

Clearly, (2) and (3) entail the negation of (1). By (3), there exist two belief systems, B and B′, where B′ is a more coherent non-trivial extension of B. Since B′ is more coherent than B and, by (2), B′ is less probable than B, (1) must be false.

How strongly supported are (2) and (3)? Claim (2) is said to require no defense: if the extension is non-trivial, then clearly "the set of beliefs containing the belief that *p* and the belief that *q* is more likely to contain only true beliefs than the set of beliefs containing both of those beliefs and, additionally, the belief that *r*" (K&W 1994, p. 130). Claim (3) is supported by the following example:

> 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 (K&W, 1994, pp. 130–131).

K&W provide the following analysis of the story. Let the original belief system contain the beliefs that (i) Dunnit had a motive; (ii) several credible witnesses report that they saw Dunnit commit the murder; (iii) a single credible witness reports that she saw Dunnit far away from the crime scene at the time of the murder. Let the extended belief system contain the same beliefs plus the additional beliefs that (iv) Dunnit has an identical twin and (v) Dunnit did it. The latter system is a non-trivial extension of the former and is more coherent than the former, which establishes (3).

K&W draw the moral that "coherence, per se, is not truth conducive" (1994, p. 132) and hence that the coherence theory of epistemic justification is untenable. Merricks (1995, p. 306) agrees with K&W that they have disproved (1), but claims that (1) does not express that coherence is truth conducive. According to Merricks, truth conduciveness should be construed as a property of individual beliefs rather than as a property of belief systems.

We do not take issue with (3), that is to say, we do not dispute the greater coherence of the extended set in comparison to the original. Clearly, the extended set hangs together better than the original and is therefore more coherent in a pre-theoretical sense. The same conclusion follows also from more theoretical considerations. One of Bonjour's celebrated coherence conditions says that "[t]he coherence of a system of beliefs is decreased in proportion to the presence of unexplained anomalies in the believed content of the system" (1985, p. 99) The detective's original belief that there is a credible witness claiming that she saw Dunnit two hundred miles away from the crime scene can plausibly be seen as confronting the detective with an unexplained anomaly. When the new evidence about the twin arrives, however, the anomaly dissolves (and no new anomalies are introduced). The extended set is thus more coherent than the original set according to Bonjour's condition.

Neither do we think that the coherentist should be willing to give up (1). As K&W correctly point out in their reply to Merricks, this form of holism is essential to the coherentist canon.

As we will see, however, (2) is much less innocent than it appears to be. It derives its prima facie plausibility from its similarity with

(2′) Any non-trivial extension of a set of propositions is less probable than the original set.

Statement (2′) is of course fully innocent, since it follows directly from the Kolmogorov axioms. The difference between (2) and (2′) is that (2′) is a claim about sets of propositions, while (2) is a claim about belief systems. What then distinguishes sets of propositions from belief systems? A belief system is not just any old set of propositions: it is a set of propositions that is believed by a particular person. Hence, whereas the probability of a set of propositions is the probability that these propositions are all true, the probability of a belief system is the probability that these propositions are all true, given that they are believed by the person in question.

In formal terms, (2′) can be expressed as follows:

(2′*) Let $S = \{p_1, \ldots, p_m\}$ and $S' = \{p_1, \ldots, p_m, p_{m+1}, \ldots p_n\}$. If S′ is a non-trivial extension of S, then $P(p_1, \ldots, p_m, p_{m+1}, \ldots p_n) < P(p_1, \ldots, p_m)$.

Let 'belpi' stand for the proposition that a particular person believes that $p_i$. Claim (2) can be expressed as follows:

(2*) Let $S = \{p_1, \ldots, p_m\}$ and $S' = \{p_1, \ldots, p_m, p_{m+1}, \ldots, p_n\}$, and let B be a belief system corresponding to S and B′ a belief system corresponding to S′. If B′ is a non-trivial extension of B, then $P(p_1, \ldots, p_m, p_{m+1}, \ldots, p_n \mid belp_1, \ldots, belp_m, belp_{m+1}, \ldots, belp_n) < P(p_1, \ldots, p_m | belp_1, \ldots, belp_m)$.

It is easy to prove that (2′*) holds true. But for K&W's argument to hold up, it is (2*) that needs to hold. We will show that (2*) is false by constructing an example in which the inequality does not hold. Throughout our argument, we adopt K&W's objective interpretation of probability. We will refer to (2*) as the Doxastic Extension Principle.

## 3. A COUNTER EXAMPLE TO THE DOXASTIC EXTENSION PRINCIPLE

Let us consider the following simple variation on the Dunnit example. Suppose that there has been a robbery. A conscientious detective would like to know whether (r) Dunnit committed the *robbery*. He consults independent witnesses to gather evidence. Although the witnesses need not be fully reliable, they all have a track record of being sufficiently reliable and our detective routinely adopts the belief that some item of evidence holds just in case there is a witness report to this effect.

Several philosophers have pointed out that we often adopt beliefs in a routine-like manner. As Isaac Levi (1991) describes the process, "[i]n routine expansion, the inquirer expands according to a program for adding new information to his state of full belief or corpus in response to external stimulation" (p. 71). A characteristic feature of routine expansion is that it does not rely on inference. While routine expansion does begin with certain assumptions or premises – namely, assumptions about the reliability of the program – the expansion adopted is not inferred from these premises (p. 74). On Levi's view, routine expansion includes consulting witnesses (p. 75). In the literature on the coherence theory, the importance of such automatically acquired, "cognitively spontaneous" beliefs has been emphasized in particular by Laurence Bonjour (1985). According to Bonjour, the distinguishing feature of these beliefs is that their occurrence at the moment in question is not the result of a discursive process of reasoning or inference (p. 129).

We will assume that each item of evidence is reported on by a single witness. After querying a bystander, the detective adopts the belief that (c) Dunnit was driving his *car* away from the crime scene at high speed. After querying one of Dunnit's neighbors, he adopts the belief that (g) Dunnit is in the possession of a *gun* of the same type as the one used in the robbery. The original belief system B contains the propositions c and g. Subsequently, a new witness steps forward: after querying the bank clerk in Dunnit's bank, the detective adopts the belief that (m) Dunnit deposited a large sum of *money* in his bank the day after the robbery. Dunnit's non-trivially extended belief system B′ now contains the additional proposition m.[4]

It is easy to construct a case in which (2*) is false on the grounds that the extended belief system B′ is *equally* probable as the original belief system B. Suppose that the witnesses are all *fully* reliable. Then

(*) $$P(\text{c, g} \mid \text{belc, belg}) = 1 = P(\text{c, g, m} \mid \text{belc, belg, belm})$$

This observation alone is sufficient to disprove the Doxastic Extension Principle (2*). But note that (2) and its formal version (2*) are actually stronger than they need to be. K&W's argument can be improved by substituting "less or equally probable" for "less probable" in (2), thereby weakening one of the premises. Hence, we need to provide a counter example against the Weak Doxastic Extension Principle that is obtained by substituting "≤ for "<" in (2*); that is, the challenge is to construct a case in which the extended belief system B′ is *more* probable than the original belief system B.

First, let us assume that there is a large number n of suspects and that each suspect stands an equal chance of having committed the robbery, so that:

(i) $P(\mathrm{r}) = 1/\mathrm{n} = 1\text{-}\mathbf{u}$, for $\mathbf{u} \approx 1$, but $\mathbf{u} \neq 1$

Second, let us assume that, although the witnesses are highly reliable, they are less than fully reliable: there is a small chance that a bystander report was forthcoming to the effect that Dunnit was speeding away from the crime scene, although he actually was not, and there is a small chance that no bystander report was forthcoming to the effect that Dunnit was speeding away from the crime scene, although he actually was. Similarly, for the two other witnesses:

(ii) $P(\mathrm{bel i|i}) = \mathbf{p}$ and $P(\mathrm{bel i|not-i}) = 1 - \mathbf{q}$ for $\mathbf{p}, \mathbf{q} \approx 1$, but $\mathbf{p}, \mathbf{q} \neq 1$ and i = c, g, m

Third, we permit probability distributions that leave a small chance that the evidence is misleading. There might be a small chance that Dunnit committed the robbery and slipped away on the subway or that Dunnit did not commit the robbery, but just happened to be speeding at the wrong time at the wrong place. Similarly, for the other items of evidence:

(iii) $P(\mathrm{i|r}) = \mathbf{s}$ and $P(\mathrm{i|not-r}) = 1 - \mathbf{t}$ for $\mathbf{s}, \mathbf{t} \approx 1$ and i = c, g, m

The next step is to introduce some assumptions of probabilistic independence. These assumptions are introduced here mainly to simplify calculations, but it is interesting to note that they characterize a common type of situation of information gathering involving independent evidence, independent witnesses and an obliquely testable hypothesis. What this means is explained below.

*Independent Evidence*. The respective items of evidence are probabilistically independent of any other items of evidence, conditional on the hypothesis. What does this mean? Suppose that we actually know whether Dunnit committed the robbery. Then there is a certain chance that he was speeding on the freeway away from the crime scene. Now suppose that we learn in addition that he is in possession of a gun of the same type as the one that was used in the robbery. Then learning this new item of evidence will not affect the chance that Dunnit was speeding on the freeway. Note that this assumption is not always fulfilled: the items of evidence may be of a nature that does not warrant this assumption. For instance, the fact that Dunnit came into the repair shop the day before the robbery for a tune-up so that his car would perform optimally at high speed, would also constitute an item of evidence that he committed the robbery, but it would not constitute an independent item of evidence.

*Independent Witnesses*. The detective's routinely acquired belief about some item of evidence is probabilistically independent of any other item of evidence or of any other of his routinely acquired beliefs, conditional on the evidence. What does this mean? Suppose that we actually know whether Dunnit was speeding on the highway. Then there is a certain chance that a reliable witness would step forward with a report to this effect and that the detective would adopt this report as a routinely acquired belief. Now suppose that we learn in addition that Dunnit was in possession of a gun of the same type as the one that was used in the robbery or that there was a witness report to this effect. Then this will not affect the chance that a reliable witness would step forward with a report to the effect that Dunnit was speeding. This assumption stipulates that each witnesses is focused on the items of evidence that he reports on and does not attend to other items of evidence or to reports about other items of evidence. What would it take for this condition not to be fulfilled? Suppose that the witnesses have been doing their own detective work: they checked out other items of evidence or talked to witnesses who reported on these items of evidence. Then their judgments of whether it was really Dunnit they saw speeding in the car could well be affected by whether there was other evidence to the effect that Dunnit committed the crime or by whether there were any reports of such evidence.

*Obliquely Testable Hypothesis*. The detective's routinely acquired beliefs about the evidence are probabilistically independent of the hypothesis, conditional on the evidence. What this means is that none of the witnesses has any direct access to whether Dunnit committed the robbery or not. This question remains hidden in the black box: the witnesses' only access to it is through the items of evidence. Suppose that we actually know that the items of evidence obtain (or that some or none obtain). Then there is a certain chance that reliable witnesses will step forward and that the detective would come to acquire beliefs to the effect that the evidence obtains. Now suppose that we learn in addition that Dunnit actually committed the crime. Then this will not affect the chance that the detective would come to acquire beliefs to the effect that the evidence obtains. What would it take for this condition not to be fulfilled? Suppose that the witnesses got a quick glimpse of the robbery scene. Then their judgments of whether it was really Dunnit who was speeding on the freeway, of whether

Dunnit really has the same gun as the one used in the crime scene, ... may be colored by what they were able to gather from the robbery scene.

We express these assumptions formally for two items of evidence, using the notation of Dawid (1979) for the propositional variables *c*, *g*, *r*, *belc*, and *belg*. (The values of the propositional variable *c* (in italics) are the proposition c and its negation not-c and similarly for other propositional variables.)

(i) $c \perp\!\!\!\perp g|r$ (Independent Evidence)

(ii) $belc \perp\!\!\!\perp g, belg|c$ and $belg \perp\!\!\!\perp c, belc|g$ (Independent Witnesses)

(iii) $belg, belc \perp\!\!\!\perp r|c, g$ (Obliquely Testable Hypothesis)

These assumptions can be readily extended for three items of evidence.

The set of conditional independence statement (i) to (iii) entails a much broader set of conditional independent statements. Dawid (1979) and Spohn (1980) introduced four properties of conditional independence statements which hold for any joint probability distribution. A set of conditional independence statements that is closed under these four properties is called a semi-graphoid. For our purposes, it will be sufficient to show that the semi-graphoid which contains (i), (ii) and (iii) also contains some additional conditional independence statements which will be helpful in our calculations (see Appendix).

We can now show that the Weak Doxastic Extension Principle is false for this particular example for some plausible values of **p**, **q**, **s**, **t** and **u**. The probability of the original belief system is:

$$(\text{Prob}_{\text{B}}) \quad P(\text{c, g} \mid \text{belc, belg})$$

The probability of the extended belief system is:

$$(\text{Prob}_{\text{B}'}) \quad P(\text{c, g, m} \mid \text{belc, belg, belm})$$

We apply Bayes' theorem to ($\text{Prob}_{\text{B}}$):

$$(\text{Prob}_{\text{B}}.1) \quad \frac{P(\text{belc, belg}|\text{c, g})P(\text{c, g})}{P(\text{belc, belg})}$$

Hence,

$$(\text{Prob}_{\text{B}}.2) \quad \frac{P(\text{belc, belg}|\text{c, g})\sum_{r} P(\text{c, g}, r)}{\sum_{c,g,r} P(\text{belc, belg}, c, g, r)}$$

where $\sum_r$ means that we are summing over all possible values of $r$, and $\sum_{c,g,r}$ we are summing over all possible combinations of values of $c$, $g$ and $r$. We apply the chain rule:

$$(\text{Prob}_\text{B}.3)\quad \frac{P(\text{belc, belg|c, g})\sum_r P(\text{c, g}|r)P(r)}{\sum_{c,g,r} P(\text{belc, belg}|c, g, r)P(c, g|r)P(r)}$$

In the appendix we have shown that the conditional independence statements that are contained in the semi-graphoid which contains (i), (ii) and (iii) permits us to do the following simplification:

$$(\text{Prob}_\text{B}.4)\quad \frac{P(\text{belc|c})P(\text{belg|g})\sum_r P(\text{c}|r)P(\text{g}|r)P(r)}{\sum_{c,g,r} P(\text{belc}|c)P(\text{belg}|g)P(c|r)P(g|r)P(r)}$$

Similarly, from ($\text{Prob}_{\text{B}'}$) we can derive that:

$$(\text{Prob}_{\text{B}'}.4)\quad \frac{P(\text{belc|c})P(\text{belg|b})P(\text{belm|m})\sum_r P(\text{c}|r)P(\text{g}|r)P(\text{m}|r)P(r)}{\sum_{c,g,m,r} P(\text{belc}|c)P(\text{belg}|g)P(\text{belm}|m)P(c|r)P(g|r)P(m|r)P(r)}$$

For definite values of the parameters **p**, **q**, **s**, **t** and **u**, we can compute $\text{Prob}_\text{B}.4$ and $\text{Prob}_{\text{B}'}.4$. Setting all parameters at .90 yields $\text{Prob}_\text{B} = P(\text{c, g, m | belc, belg, belm}) = 0.8910 > 0.7562 = P(\text{c, g | belc, belg}) = \text{Prob}_{\text{B}'}$, which is precisely the desired result.

How is it that the level of the reliability of the witnesses and the coherence of the extended belief system brings about an increase in the chance that the content of the extended belief system is true in our example? There is an increase in the chance that the content of the extended belief system is true due to the level of the reliability of the witnesses and the coherence of the extended belief system. As to the reliability of the witnesses, we have chosen the parameters **p** and **q** so that the witnesses are highly but not fully reliable. As to coherence, it is notoriously difficult to construct a precise coherence measure of how well a belief set fits together. But the following observation should be uncontroversial. Take the prior probability of each genuine subset $S$ of {c, g, m} and the conditional probability of each genuine subset $S$ of {c, g, m} on the remaining elements of {c, g, m}.

Certainly, {c, g, m} is a highly coherent set, when the conditional probability strongly exceeds the prior probability of each S. The prior probability of S depends on **s**, **t** and **u**, and only on these parameters. For example, $P(\text{c, g}) = \sum_r P(\text{c}|r)P(\text{g}|r)P(r) = \mathbf{s}^2(1-\mathbf{u}) + (1-\mathbf{t})^2\mathbf{u}$. Likewise, the conditional probability of S on the remaining elements depends on **s**, **t** and **u** and only on these parameters. For example,

$$P(\text{c, g} \mid \text{m}) = \frac{\sum_r P(\text{c}|r)P(\text{g}|r)P(\text{m}|r)P(r)}{\sum_r P(\text{m}|r)P(r))} = \frac{\mathbf{s}^3(1-\mathbf{u}) + (1-\mathbf{t})^3\mathbf{u}}{\mathbf{s}(1-\mathbf{u}) + (1-\mathbf{t})\mathbf{u}}.$$

We have chosen these parameters so that {c, g, m} qualifies as a highly coherent set. For each singleton $S$, $P(\text{i}|\text{j, k}) = 0.82 \gg 0.18 = P(\text{i})$ and for each pair $S$, $P(\text{i, j}|\text{k}) = 0.41 \gg 0.09 = P(\text{i, j})$, for i, j, k = c, g, m and $\text{i} \neq \text{j}$, $\text{j} \neq \text{k}$, $\text{i} \neq \text{k}$. This probabilistic feature of the belief set is certainly sufficient to warrant the assessment that the belief set in question is highly coherent.[5]

## 4. DISCUSSION

Klein and Warfield's argument against the truth conduciveness of coherence is based on the premise that a belief system is always (except in trivial cases) less likely to be true as a whole than any proper subsystem. We have seen that this premise is false: a larger belief system need not be less likely to be true but can even be more likely to be true than its parts. Hence, K&W's argument fails. The general lesson is that by believing more things you do not automatically run a greater risk of being in error, and so you do not thereby necessarily come in conflict with the epistemic goal to shun error.

In asking whether a belief system is likely to be true, we consider the probability of the propositions in the belief system, conditional on having acquired the beliefs in some non-inferential manner. The coherentist is best understood as advancing at least the following thesis: Coherence is truth conducive in the sense that if a set of propositions $\{p_1, \ldots, p_n\}$ is more coherent than a set $\{q_1, \ldots, q_n\}$, then $P(p_1, \ldots, p_n|\text{bel}p_1, \ldots, \text{bel}p_n) > P(q_1, \ldots, q_n|\text{bel}q_1, \ldots, \text{bel}q_n)$. In other words, the joint probability of a more coherent set of propositions, given that it is the content of a belief system of cognitively spontaneous beliefs, is greater than the joint probability of a less coherent set of propositions, given that it is the content of a belief system of cognitively spontaneous beliefs.[6]

But this cannot be the whole story about the connection between coherence and truth. It is implausible to think that coherence has the property of truth conduciveness, unless we impose certain constraints on the acquisition of the beliefs. First it must be assumed that the beliefs are acquired through processes that are to some extent reliable. As C.I. Lewis (1946, pp. 338–62) emphasized, the coherence of beliefs that are acquired through a fully unreliable process will do nothing for their credibility. Second, as was equally clear to Lewis, the beliefs must be, at least to some degree, independently acquired. Following Lewis, Bonjour (1985, p. 148) writes that "so long as cognitively spontaneous beliefs are genuinely independent of each other, their agreement will eventually generate credibility". In addition, to assess the impact of coherence, we also need to build in a *ceteris paribus* clause. Suppose that one set of beliefs is more coherent than another. Yet because the beliefs in the more coherent set are obtained through less reliable processes or because there was more dependence between the witness reports, the posterior joint probability of the less coherent set turns out to be higher. Should we take such a scenario to present a counterexample to the truth conduciveness of coherence? Presumably not. The coherentist claim is that a more coherent set of propositions is more likely to be true assuming that all other factors that affect the posterior probability are kept constant. We leave it as an open question here what factors need to be kept fixed and under what conditions coherence indeed turns out to be truth conducive. These issues are discussed at length in Bovens and Olsson (2000), Olsson (2001) and Olsson (2002).

We have followed K&W in assuming the probabilities to be objective. We have also made certain assumptions concerning the reliability of the witnesses. Hence, one might object that our argument against K&W is really of an 'externalist' or a 'reliabililist' nature, and reliabilism and coherentism are commonly seen as opposites on the epistemological scale. It was never our intention to defend a *bona fide* coherentism, however one may wish to conceive of such a doctrine. The view which emerges from this study is that reliabilism and coherentism are not opposites but rather complements in the following sense: The reliability of the process of belief acquisition and the coherence of our beliefs system are both factors that influence the posterior probability of the content of our belief system. No headway is being made by overemphasizing one factor at the expense of the other. A reasonable theory of knowledge and justification must be ecumenical and ascribe a proper role to both factors.

We conclude that, under certain independence assumptions, if we gain information from highly but not fully reliable witnesses about highly coherent sets of propositions, then our extended belief system will be more probable than our original belief system and we will have a counter example to the (Weak) Doxastic Extension Principle. This opens up an interesting area of research: what is the precise relation between the coherence of a belief system, the reliability of the information sources and the joint probability of the original and the extended belief system? A more complete investigation into this matter will have to await another occasion.

## 5. APPENDIX

There are a number of sets of conditional independence statements that permits us to do this simplification. One such set is the set containing: (a) $belc \perp\!\!\!\perp belg|c, g$; (b) $belc \perp\!\!\!\perp g|c$; (c) $belg \perp\!\!\!\perp c|g$; (d) $c \perp\!\!\!\perp g|r$; (e) $belg, belgc \perp\!\!\!\perp r|c, g$. We need to show (I) that (a) through (e) are contained in the semi-graphoid which is the closure of {(i) Independent Evidence, (ii) Independent Witnesses, (iii) Obliquely Testable Hypothesis} under the semi-graphoid properties, i.e., symmetry, decomposition, weak union and contraction; and (II) that (a) through (e) permit us to derive ($\text{Prob}_\text{B}$.4) from ($\text{Prob}_\text{B}$.3). The proof in (I) only requires applications of the properties of decomposition and weak union. For any set of variables $\alpha$, $\beta$, $\gamma$, and $\delta$, decomposition states that, if $\alpha \perp\!\!\!\perp \beta, \gamma|\delta$, then $\alpha \perp\!\!\!\perp \beta|\delta$ and weak union states that, if $\alpha \perp\!\!\!\perp \beta, \gamma|\delta$, then $\alpha \perp\!\!\!\perp \beta|\gamma, \delta$. For a clear statement of the semi-graphoid properties, see Pearl (1988) pp. 117–118 and Castillo, E. (*et. al.*) (1997) pp. 180–195.

I. (d) is identical to (i) and (e) is identical to (iii); (b) and (c) follow from (ii) by decomposition; (a) follows from (ii) by weak union.

II. Assume ($\text{Prob}_\text{B}$.3). Successively apply (a), (b) and (c) to the first factor in the numerator. Apply (d) to the first factor of the sum in the numerator. Successively apply (e), (a), (b) and (c) to the first factor in the sum of the denominator. Apply (d) to the second factor of the sum in the denominator. The result of these operations is ($\text{Prob}_{\text{B}.4}$).

## ACKNOWLEDGEMENTS

Thanks to the two anonymous referees of this journal for helpful comments and suggestions. Luc Bovens's research was supported by grants of the National Science Foundation (Science and Technologies Studies – SES 00-80580) and by the Alexander von Humboldt Foundation, the Federal

Ministry of Education and Research, and the Program for Investment in the Future (ZIP) of the German Government. Erik J. Olsson's research was supported by the German Research Council (*Deutsche Forschungsgemeinschaft*) as a contribution to the project Logic in Philosophy (*Logik in der Philosophie*).

## NOTES

[1] Levi (1991), p. 81.

[2] For a more thorough examination of some other aspects of the connection between coherentism, conditional independence structures and Bayesian networks, see Bovens and Olsson (2000).

[3] See Bonjour (1985), chapter 8.

[4] One might object that we could have discussed an extension from one to two beliefs, rather than from two to three beliefs. From a mathematical point of view, this would indeed have been sufficient for the purpose of rejecting (2*). However, from an epistemological point of view, the one-proposition case is problematic. The reason is that we want to compare two belief systems with respect to their coherence and one belief hardly qualifies as a belief *system*. Moverover, Nicolas Rescher has argued that coherence is a concept that simply does not apply to singletons: "Coherence is . . . a feature that propositions cannot have in isolation but only in groups containing several – i.e., at least two – propositions" (1973, p. 32). Hence, we cannot compare a singleton with a set of two or more propositions with respect to coherence. The relation "more coherent than" is undefined if one of the relata is a singleton. In order to avoid such conceptual problems, we consider an extension of a belief system from two to three propositions.

[5] For a general probabilistic criterion of belief expansion see Bovens and Hartmann (2000), and for a measure of coherence that imposes a partial ordering on a set of belief systems, see Bovens and Hartmann (forthcoming).

[6] In an attempt to capture the sense of truth-conduciveness underlying the epistemology of Bonjour, Cross (1999) spells out a conditional account that has certain commonalities with our approach. For Cross, truth-conduciveness is a function of three parameters, viz. the coherence of the belief system, the stability of the belief system and the length of the subject's belief history. Cross's approach may be accurate as an explication of Bonjour's intentions, though he does not (nor does he intend to) lay out how it is that the particular aspect of coherence is truth-conducive.

## REFERENCES

Bonjour, L.: 1985, *The Structure of Empirical Knowledge*, Harvard University Press, Cambridge, MA.

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

Bovens, L. and S. Hartmann: 2000, 'Coherence, Belief Expansion and Bayesian Networks', in C. Baral and M. Truszczynski (eds), *Proceedings of the*

*8th International Workshop on Non-Monotonic Reasoning, NMR 2000*. (http://xxx.lanl.gov/abs/cs.AI/0003041).

Bovens, L. and S. Hartmann: forthcoming, 'Solving the Riddle of Coherence'.

Cross, C. B.: 1999, 'Coherence and Truth-Conducive Justification', *Analysis* **59**, 186–193.

Dawid, A. P.: 1979, 'Conditional Independence in Statistical Theory', *J. Royal Statist. Soc.*, ser. B 41 (No. 1), pp. 1–31.

James, W.: 1897, *The Will to Believe and Other Essays*, Longman, New York.

Jensen, F. V.: 1996, *An Introduction to Bayesian Networks*, University College London Press, London.

Klein, P. and T. A. Warfield: 1994, 'What Price Coherence?', *Analysis* **54**, 129–132.

Klein, P. and T. A. Warfield: 1996, 'No Help For the Coherentist', *Analysis* **56**(2), 118–121.

Levi, I.: 1991, *The Fixation of Belief and its Undoing*, Cambridge University Press, Cambridge.

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

Merricks, T.: 1995, 'On Behalf of the Coherentist', *Analysis* **55**(4), 306–309.

Neapolitan, R. E.: 1990, *Probabilistic Reasoning in Expert Systems*, John Wiley & Sons, New York, NY.

Olsson, E. J.: 2001, 'Why Coherence is not Truth-Conducive', *Analysis* **61**(3), 236–241.

Olsson, E. J.: 2002, 'What is the Problem of Coherence and Truth?', *The Journal of Philosophy* **XCIX**(5), 246–272.

Pearl, J.: 1988, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Francisco, CA.

Pearl, J.: 2000, *Causality*, Cambridge University Press, Cambridge.

Rescher, N.: 1982 [1973], *The Coherence Theory of Truth*, The University Press of America, Washington.

Spohn, W.: 1980, 'Stochastic Independence, Causal Independence, and Shieldability', *Journal of Philosophical Logic* **9**, 73–99.

Luc Bovens
University of Colorado at Boulder
Department of Philosophy
CB 232
Boulder CO 80309
U.S.A.

Erik J. Olsson
Universität Konstanz
Fachgruppe Philosophie
Postfach 5560 D21
78457 Konstanz
Germany