LUND UNIVERSITY

Bayesian Epistemology

Olsson, Erik J

Published in:
Introduction to Formal Philosophy

2018

Document Version:
Peer reviewed version (aka post-print)

Bayesian Epistemology

Erik J. Olsson

Abstract: Bayesian epistemology provides a formal framework within which concepts in traditional epistemology, in particular concepts relating to the justification of our beliefs, can be given precise definitions in terms of probability. The Bayesian approach has contributed clarity and precision to a number of traditional issues. A salient example is the recent embedding of the so-called coherentist theory of epistemic justification in a Bayesian framework shedding light on the relation between coherence and truth as well as on the concept of coherence itself. Starting with the early work of Condorcet, the calculus of probability has proved to be a useful tool in the study of social aspects of knowledge as it is pursued in social epistemology.

1. Two problems of probabilistic coherence

Let us start by examining the two concepts involved in the term "Bayesian epistemology".

First, we have the term "Bayesian" which in this context denotes a plethora of theories and approaches that make use of probability in the elucidation of phenomena having to do with our beliefs about the world. One aspect of the Bayesian approach, also called Bayesianism, is the representation of a state of belief as an assignment of probabilities to a set of propositions. Typically, Bayesians feel uncomfortable in assigning any proposition probability 0 or 1. Rather, they recommend assigning probabilities strictly between 0 and 1. One reason for this is the so-called betting interpretation of probabilities according to which assigning a probability means that you are willing to accept certain bets. Assigning probability 1 to a proposition means then that you are willing to bet everything – your life, your family etc. – on p being true. Since we are rarely willing to

bet everything on a given (empirical) proposition being true, we should avoid assigning probability 1 to any such proposition, the Bayesian concludes.

This is the static aspect of Bayesianism. There is also a dynamic aspect enshrined in the recommendation that a rational inquirer should update her beliefs by conditionalizing on the new evidence in accordance with Bayes' rule. Let us first define the conditional probability of the hypothesis h given evidence e: $P(h \mid e) = P(e \mid h)P(h)/P(e)$. This equation still does not state anything about how the inquirer's probabilities should change given new evidence. Bayes' rule, sometimes also referred to as the principle of conditionalization, states that the inquirer should, upon receiving new evidence e, update her probability in h so that the latter corresponds to the conditional probability of h given e. In other words, $P^*(h) = P(h \mid e)$, where $P^*(h)$ is the new probability of h given evidence e. These two fundamental assumptions of Bayesianism have inspired a huge debate in philosophy of science and statistical theory, as well as in economics and decision theory. The reader is referred to Talbott (2016) for an overview.

Let us proceed now to the term "epistemology" or "theory of knowledge". Epistemology is concerned with various aspects of knowledge. What is the nature of knowledge and how should the concept be defined? What sources give rise to knowledge? How far does our knowledge extend – are there limits in principle? Do we have knowledge at all – or do we have to accept some form of skepticism? If we know, do we know that we do? And so on. Traditionally, the answer to the first question – about the nature of knowledge – has been knowledge amounts to justified true belief. If you have a belief and you entertain that belief with justification, then you know, provided of course that the proposition in question is true. Believing means in this context being sure or fully convinced of the truth of the proposition.

These standard characterizations of Bayesianism and epistemology reveal that it is not unproblematic to coherently combine the two into "Bayesian epistemology". Just to raise one question: How does the fact that knowledge requires full conviction square with the Bayesian recommendation not to assign 1 to any given empirical proposition? Does not skepticism about empirical knowledge ensue? Perhaps unsurprisingly, recent texts on "Bayesian epistemology" in fact do not address in any great detail problems in traditional epistemology but rather uses the term as roughly synonymous with "Bayesianism" (e.g.

Talbott, 2016). Even though there are certain tensions to be overcome, a Bayesian approach can in fact be very effective in the elucidation of the justification part of the traditional concept of knowledge. This holds in particular of so-called coherentist accounts of justification.

The truth conduciveness of coherence. Pre-systematically, coherence is a good thing. If a set of beliefs is coherent, we tend to think that it is plausibly true, and that a more coherent set is more likely to be true than a less coherent one. Consider however the following example from Klein and Warfield (1994, 130-131):

> 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 beliefs 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.

Let the original belief system of the detective contain the beliefs that (1) Dunnit had a motive; (2) several credible witnesses report that they saw Dunnit commit the murder; (3) 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 believes plus the additional beliefs that (4) Dunnit has an identical twin and (5) Dunnit did it. Then we would say that the extended system is more coherent than the original belief system. So we should expect the former to be more likely to be true. However, the extended system contains more propositions than the original system, and hence the probability of the conjunction of the propositions in the extended system must be lower than the probability of the conjunction of the propositions in the original system: disregarding some trivial special cases, the probability of a bigger conjunction is lower than the probability of a smaller conjunction. So, despite being more coherent, the extended system is actually less likely to be true. So, coherence is after all not correlated with plausible truth.

Defining coherence. There have been few convincing proposals for how to define coherence in traditional epistemology. The attempt to spell our coherence in purely logical terms, e.g. by A. C. Ewing (1934), was soon seen to be too restrictive. Most other proposals suffer from serious incompleteness or imprecision. A case in point is the account due to Laurence BonJour (1985), who regards coherence to be a concept with a multitude of different aspects, corresponding to the following *coherence criteria* (ibid.: 97-99):

(1) A system of beliefs is coherent only if it is logically consistent.

(2) A system of beliefs is coherent in proportion to its degree of probabilistic consistency.

(3) The coherence of a system of beliefs is increased by the presence of inferential connections between its component beliefs and increased in proportion to the number and strength of such connections.

(4) The coherence of a system of beliefs is diminished to the extent to which it is divided into subsystems of beliefs which are relatively unconnected to each other by inferential connections.

(5) The coherence of a system of beliefs is decreased in proportion to the presence of unexplained anomalies in the believed content of the system.

Now it could well happen that one system S is more coherent than another system T in one respect, whereas T is more coherent than S in another. Perhaps S contains more inferential connections than T, which is less anomalous than S. If so, which system is more coherent in an overall sense? A difficulty pertaining to theories of coherence that construe coherence as a multifaceted concept is to specify how the different aspects are to be amalgamated into one overall coherence judgment. Bonjour's theory remains silent on this important point and, as we shall see, in several other regards as well.

2. A Bayesian analysis of the Dunnit example

Let us state the argument in more precise terms. A central claim in the coherence theory which has strong intuitive backing is the following:

(A) The more coherent a set is, the more probable it is.

Let us say that an extension $K´$ of a set K is non-trivial if some of the beliefs that are $K´$ but not in K neither follow logically from K, nor have a probability of 1. Klein and Warfield's argument against (A) rests on the following premises:

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

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

But, so the argument goes, (B) and (C) taken together contradict (A).

Let us look at the support for (B) and (C). While (B) is taken for granted, (C) is supported by the above Dunnit example. It is difficult to question (C). Intuitively the members of the extended set in the Dunnit example hangs better together than the elements of the original set. Also, the original set contains an anomaly which is resolved through the introduction of the beliefs that Dunnit did it and has an identical twin who the witness providing the alibi mistook for Dunnit. Because no new anomaly is thereby introduced, it follows from Bonjour's fifth criterion that the extended set is more coherent.

But what about (B)? It derives support from its similarity with

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

That claim follows directly from the laws of probability and is therefore entirely innocent. But notice that (B´) is about sets of propositions, whereas (B) is about belief systems. What is the difference? A belief system is not any old set of propositions but a

set of propositions believed to be true by a subject. 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 true, given that they are believed by the person in question. The former is an unconditional and the latter a conditional probability.

Let $S = \{p_1, \ldots, p_m\}$ and $S´ = \{p_1, \ldots, p_m, p_{m+1}, \ldots, p_n\}$. Moreover, let B be a belief system corresponding to S and $B´$ be a belief system corresponding to $S´$. Formally, (B´) can be expressed as follows:

(B´*) If $S´$ is a non-trivial extension of S, then $P(p_1, \ldots, p_m, p_{m+1}, \ldots, p_n) < P(p_1, \ldots, p_m)$.

The claim (B) should rather be understood as follows:

(B*) If $B´$ is a non-trivial extension of B, then $P(p_1, \ldots, p_m, p_{m+1}, \ldots, p_n \mid belp_1, \ldots, belp_m, belp_{m+1}, \ldots, belp_n) < P(p_1, \ldots, p_m \mid belp_1, \ldots, belp_n)$,

where $belp_i$ states that the subject believes proposition p_i.

For the Dunnit argument it is (B*) that needs to hold, not (B´*). It can be shown however that (B*) is false. There can be non-trivial extensions of a belief system that are more probable than the original belief system. Suppose again that a robbery has been committed. A detective wishing to find out whether Dunnit did it (call that proposition r) consults independent witnesses that have a track-record of being sufficiently reliable so that the detective can routinely trust their reports. This reminds us of Bonjour's "cognitively spontaneous beliefs" which play a crucial role in his epistemology. We assume that the detective believes something just in case a witness has said so.

Suppose that the first witness reports that Dunnit was driving his car away from the crime scene at high speed (c) and the second that Dunnit is in the possession of a gun of the relevant type (g). The original belief system contains the propositions c and g. Now a new witness steps forward, claiming that Dunnit deposited a large sum of money in his bank the day after the robbery (m). The extended belief system contains the propositions c, g and m. The key notions of reliability and witness independence can be expressed in

probability theory. For instance, that a given witness is a reliable belief producer can be expressed as follows:

$P(\text{beli} \mid i) = \mathbf{p}$ and $P(\text{beli} \mid \text{not-i}) = 1\text{-}\mathbf{q}$ for $\mathbf{p}, \mathbf{q} \approx 1$ and $i = c, g, m$.

Hence, the probability that you form the belief, if it is true, should be high, and the probability that you form the belief, if it is false should be low.

That the beliefs are independently held means that they there is no direct influence between the testimonies upon which they were based. This can be captured by saying that the detective's routinely acquired belief about some item of evidence is probabilistically independent of any other item of evidence or any other of his routinely acquired beliefs, conditional on the that item of evidence. We express this 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 are the propositions c and its negation not-c and similarly for the other propositional variables.)

belc $\perp$ *g*, *belg* $\mid$ *c* and *belg* $\perp$ *c*, *belc* $\mid$ *g*

The first part of this statement is read *belc* is independent of *g* and *belg* given *c*, which is sometimes expressed by saying that *c* "screens off" *belc* from *g* and *belg*. This implies for instance that belc is independent of not-g and belg given not-c.

With a few additional assumptions it can now be proved that the extended belief system is more probable than the original system:

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

For more details and a proof, see Bovens and Olsson (2002).

3. Bayesian accounts of coherence

7

Let us now return to the problem of how to define coherence. Bonjour's account serves to illustrate another general difficulty. The third criterion stipulates that the degree of coherence increases with the number of inferential connections between different parts of the system. As a system grows larger the probability is increased that there will be relatively many inferentially connected beliefs. For a smaller system, this is less likely. Hence, there will be a positive correlation between system size and the number of inferential connections. Taken literally, Bonjour's third criterion implies, therefore, that there will be a positive correlation between system size and degree of coherence. But this is not obviously correct.

Here is another general challenge for those wishing to give a clear-cut account of coherence. Suppose a number of eye witnesses are being questioned separately concerning a robbery that has recently taken place. The first two witnesses, Robert and Mary, give exactly the same detailed description of the robber as a red-headed man in his forties of normal height wearing a blue leather jacket and green shoes. The next two witnesses, Steve and Karen, also tell exactly the same story but only succeed in giving a very general description of the robber as a man wearing a blue leather jacket. So here we have two cases of exact agreement. In one case, the agreement concerns something very specific and detailed, while in the other case it concerns a more general proposition. This raises the question of which pair of reports is more coherent. Should we say that agreement on something specific gives rise to a higher degree of coherence, perhaps because such agreement seems more "striking"? Or should we rather maintain that the degree of coherence is the same, regardless of the specificity of the thing agreed upon?

The rich literature on Bayesian coherence measures provides various answers to these questions. Here are the two most discussed measures:

$$C_1(p_1,\ldots,p_n) = P(p_1 \wedge \ldots \wedge p_n)/P(p_1) \times \ldots \times P(p_n)$$

$$C_2(p_1,\ldots,p_n) = P(p_1 \wedge \ldots \wedge p_n)/P(p_1 \vee \ldots \vee p_n)$$

C_1 was put forward in Shogenji (1999) while C_2 was tentatively proposed in Olsson (2002) and, independently, in Glass (2002). As the reader can verify, C_1 is sensitive to

size as well as to specificity, while this is not so for C_2. It has been suggested, therefore, that these two measures actually measure two different things. While C_2 captures the degree of agreement of the proposition in a set, C_1 is more plausible as a measure of how *striking* the agreement is. See Olsson (2002) and also Bovens and Olsson (2000) for a discussion of agreement vs. striking agreement. Since the appearance of these two measure, a large number of other alternative measures have been proposed, many of which are considered in Olsson and Schubert (2007).

One influential thought in traditional epistemology is that coherence is somehow linked with "mutual support". The Bayesian way of thinking of support is in terms of a confirmation measure. Douven and Meijs (2007) have proposed a general scheme for defining coherence measures given a measure S of degree of confirmation. For two propositions p and q, their suggestion takes the following form:

$$C_3(p,q) = ½ \, (S(p,q) + S(q,p))$$

Thus, the degree of coherence of a set of two propositions depends on how much they confirm each other on the average. In order to turn this scheme into a definite measure of coherence, we have to specify a particular measure of confirmation, of which there is no shortage in the Bayesian literature. Douven and Meijs's preferred choice is the difference measure advocated by Gillies (1986) and others:

$$C_4(p,q) = P(p|q) - P(p)$$

Plugging in this measure in Douven and Meijs's recipe yields the following formula:

$$C_5(p,q) = ½ \, (P(p|q) - P(p) + P(q|p) - P(q))$$

But there are of course a whole range of other confirmation measures that could just as well have been employed, e.g., the ratio measure preferred by Schlesinger (1995) and others:

$C_6(\mathrm{p,q}) = P(\mathrm{p|q})/P(\mathrm{p})$

As is easily seen, the ratio measure of confirmation coincides with the Shogenji measure of coherence for the case of two propositions.

4. Impossibility results for coherence and truth

The paper by Klein and Warfield and also Michael Huemer (1997) spurred an intense debate on the relation between coherence and truth or high probability, a debate which is still on-going. The most thought-provoking results concern the possibility of finding a measure of coherence that is *truth conducive* in the following sense: if a set of beliefs *A* is more coherent than another set of beliefs *B*, then the probability of *A* is higher than the probability of *B*. Here it is assumed that the beliefs in question are somewhat reliable and independently held. Finding such a measure was first stated as an open problem in Olsson (2002). An impossibility result to that effect was first proved by Luc Bovens and Stephan Hartmann in their 2003 book. A different impossibility theorem was proved in Olsson (2005).

These impossibility results give rise to a mind-boggling paradox. How can it be that we trust and rely on coherence reasoning, in everyday life and in science, when in fact coherence is not truth conducive? Since the impossibility results were published a number of proposals have been made for how to avoid the anomaly they present us with. Olsson and Schubert (2007) observed that, while coherence falls short of being truth conducive, it can still be "reliability conducive," i.e. more coherence, according to some measures, entail a higher probability that the sources are reliable, at least in a paradigmatic case. For a further development of this idea, see Schubert (2011). Staffan Angere (2007, 2008) has argued, based on the results of computer simulations, that the fact that coherence fails to be truth conducive in the sense just referred to does not prevent it from being connected with truth in a weaker, defeasible sense: almost all coherence measures that have an independent standing in the literature satisfy the condition that *most* cases of higher coherence are also cases of higher likelihood. Other researchers have proposed other ways of reconciling the impossibility results with our

ordinary reliance on coherence. For an up-to-date overview of the debate, see Olsson (2017).

5. Bayesian social epistemology

Following C. I. Lewis (1946), most Bayesian coherence theorists take as their paradigm case a scenario involving a number of witnesses giving coherent testimonies. This is then taken to be analogous to the situation upon which traditional coherence theorists have been most interested: the coherence of one person's beliefs. It is perfectly possible to by-pass the second issue so as to focus only on witness scenarios, in which case the study falls under the area known as social epistemology. Bovens and Hartmann (2003) elaborate on witness coherence and their book contains further references. A closely related topic is the Bayesian study of voting and the famous Condorcet Jury Theorem which states, roughly, that if voters are independent and somewhat reliable, the majority is more likely to have the right answer than anyone in the minority. Moreover, the chance that the majority is right approaches 1 as more voters are added. See for instance Goodin and List (2001) for more on this.

The Jury Theorem belongs, more generally, to what Alvin I. Goldman (1999) calls veritistic social epistemology which aims to evaluate social practices, jury voting being but one case, in terms of their veritistic outputs, where veritistic outputs includes states like knowledge, error and ignorance. Goldman focuses on the tendency of practices to produce *true belief* in the participants, true belief representing in his view a weak form of knowledge. Thus, states of true belief have *fundamental* veritistic value or disvalue, whereas practices have *instrumental* veritistic value insofar as they promote or impede the acquisition of fundamental veritistic value.

Let us now turn to the very concept of veritistic value. Goldman's main proposal is that degrees of belief (DB) have veritistic value relative to a question Q, so that any DB in the true answer to Q has the same amount of V-value as the strength of the DB. Goldman represents strength of belief as subjective probability. In Goldman's terminology, V-value of $DB_x(true) = \mathbf{x}$. Suppose, for example, that Mary is interested in the question whether it will rain tomorrow. If the strength of Mary's belief that it will rain

tomorrow is .8, and it will in fact rain tomorrow, then the V-value of Mary's state of belief vis-à-vis the rain issue is .8.

Suppose that a question begins to interest agent S at time t_1, and S applies a certain practice π in order to answer the question. The practice might consist, for instance, in a certain perceptual investigation or in asking a friend. If the result of applying π is to increase the V-value of the belief states from t_1 to t_2, then π deserves positive credit. If it lowers the V-value it deserves negative credit. If it does neither, it is neutral with respect to instrumental V-value. There is more complexity to come, however. In evaluating the V-value of a *practice*, we usually cannot focus merely on the one agent scenario. As Goldman notes, "[m]any social practices aim to disseminate information to multiple agents, and their success should be judged by their propensity to increase the V-value of many agents' belief states, not just the belief states of a single agent" (1999, 93). This is why we should be interested in the *aggregate* level of knowledge, or true belief, of an entire community (or a subset thereof).

Consider a small community of four agents: S_1-S_4. Suppose that the question of interest is whether p or not-p is true, and that p is in fact true. At time t_1, the several agents have DBs vis-à-vis p as shown in the corresponding column (see Table 1). Practice π is then applied, with the result that the agents acquire new DBs vis-à-vis P at t_2 as shown in the column under t_2.

	t_1	t_2
S_1	DB(p) = .40	DB(p) = . 70
S_2	DB(p) = .70	DB(p) = .90
S_3	DB(p) = .90	DB(p) = .60
S_4	DB(p) = .20	DB(p) = .80

Table 1

At t_1 the group's mean DB in p is .55, so that .55 is their aggregate V-value at t_1. At t_2, the group's mean DB in p is .75, so that this is their new aggregate V-value. Thus the group displays an increase of .20 in its aggregate V-value. Hence the practice π displays positive V-value in this application.

A further issue is that there is a need to consider not just one application of a practice but many such applications. In evaluating a practice, we are interested in its performance across a wide range of applications. In order to determine the V-value of the practice π in our example we would have to study how well it fares in other applications as well. This would presumably mean, among other things, varying the size of the population of inquirers as well as allowing it to operate on other initial degrees of belief. Once we have isolated the relevant set of applications against which the practice is to be measured, we can take its average performance as a measure of its V-value.

It follows from these considerations that, when assessing the V-value of a practice, we need to "average" twice. For each application A_i of the practice, we need to assess the average effect E_i it had on the degrees of belief of the members of the society. The V-value of the practice is then computed as the average over all the E_is.

As one can imagine, the task to compute the V-value of a social practice can become quite complicated in practice. For that reason, researchers have been interested in delegating it to computers. See Olsson (2011) for a description of the simulation framework Laputa which allows V-values to be computed automatically for a variety of social practices.

6. The value of Bayesian epistemology

Pursuing Bayesian epistemology, as understood here and arguably in Bovens and Hartmann (2003), means translating concepts and ideas from epistemology into the language of probability, especially concepts that relate to the way in which our beliefs are justified. This brings with it a number of advantages, many of which pertain to the use of formal methods generally. One has already been made: by means of formalization vague

13

or ambiguous concepts can be made precise and different senses distinguished. This was amply illustrated in our discussion of various ways of defining the concept of coherence – the central concept in the coherentist theory of justification – in probabilistic terms. Further, once a problem has been translated into probability theory, it can be handled in a more objective fashion than was previously possible. Our Bayesian treatment of the Dunnit example due to Klein and Warfield illustrates this advantage allowing it to be rigorously proved that one of their premises is false. The same example pinpoints another virtue of formalization: the possibility of making and upholding delicate distinctions that are difficult to express and sustain in ordinary language, i.e., the distinction between any old propositions and propositions that are believed to be true by some inquirer, and the implications of that difference for the probability of a set. See Hansson (2000) for an illuminating discussion of the value of formalization.

Finally, formalization in a standard formal framework, probability being no exception, furthers the important scientific virtues of unity and integration. Thus, the marriage between coherence and probability has led to a tighter connection between epistemology and other areas of philosophy and science in which probability plays a major role. As we saw, authors have explored the rather obvious connection to confirmation theory, including Branden Fitelson (2003). Links to artificial intelligence – Bayesian networks and fuzzy logic respectively – are established in Bovens and Olsson (2000) and Glass (2006).

References and proposed readings

Angere, S. (2007), "The Defeasible Nature of Coherentist Justification", *Synthese* 157(3): 321-335.
Angere, S. (2008), "Coherence as a Heuristic", *Mind* 117(465): 1-26.
Bonjour, L. (1985). *The Structure of Empirical Knowledge*, Cambridge, Mass.: Harvard University Press.
PROPOSED READING Bovens, L, and Hartmann, S. (2003), *Bayesian Epistemology*, Clarendon Press: Oxford. [A technically competent survey of applications of

Bayesianism to problems in informal philosophy, including an impossibility result for coherence.]

Bovens, L. and Olsson, E. J. (2000), "Coherentism, Reliability and Bayesian Networks", *Mind* 109: 685-719.

Bovens, L. and Olsson, E. J. (2002), "Believing More, Risking Less: On Coherence, Truth and Non-trivial Extensions", *Erkenntnis* 57: 137-50.

David, A. P. (1979), "Conditional Independence in Statistical Theory", *Journal of the Royal Statistical Society*, ser. B 41(1): 1-31.

Douven, I., and Meijs, W. (2007), "Measuring Coherence", *Synthese* 156(3): 405-425.

Ewing, A. C. (1934), *Idealism: A Critical Survey*, Methuen: London.

Fitelson, B. (2003), "A Probabilistic Measure of Coherence", *Analysis* 63: 194-199.

Gillies, D. (1986), "In Defense of the Popper-Miller Argument", *Philosophy of Science* 53: 110-113.

Glass, D. H. (2002), "Coherence, Explanation and Bayesian Networks", in M . O'Neill and R. F. E. Sutcliffe et al (Eds.) *Artificial Intelligence and Cognitive Science*, Berlin: Springer-Verlag, Lecture Notes in Artificial Intelligence 2464, 177-182.

Goldman, A. I. (1999), *Knowledge in a Social World*, Oxford University Press.

Goodin, R. E, and List, C. (2001), "Epistemic Democracy: Generalizing the Condorcet Jury Theorem", *Journal of Political Philosophy* 9(3): 277-306.

Glass, D. H. (2006), "Coherence Measures and their Relation to Fuzzy Similarity and Inconsistency in Knowledge Bases", *Artificial Intelligence Review* 26(3): 227-249.

Hansson, S. O. (2000), "Formalization in Philosophy", *Bulletin of Symbolic Logic* 6(2): 162-175.

Huemer, M. (1997), "Probability and Coherence Justification", *Southern Journal of Philosophy* 35: 463-72.

Klein, P. and Warfield, T. A. (1994), "What Price Coherence?", *Analysis* 54: 129-32.

PROPOSED READING Lewis, C. I. (1946), *An Analysis of Knowledge and Valuation*, LaSalle: Open Court. [An early classic on probabilistic coherence theory.]

Olsson, E. J. (2002), "What is the Problem of Coherence and Truth?", *The Journal of Philosophy* 99: 246-72.

PROPOSED READING Olsson, E. J. (2005), *Against Coherence: Truth, Probability, and Justification*, Oxford University Press. [Applies Bayesian concepts to informal coherence theory, including an impossibility theorem.]

Olsson, E. J. (2011), "A Simulation Approach to Veritistic Social Epistemology", *Episteme* 8(2): 127-143.

PROPOSED READING Olsson, E. J. (2017), "Coherentist Theories of Epistemic Justification", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/justep-coherence/>. [A regularly updated overview over the debate over the coherence theory of justification with particular emphasis on Bayesian approaches.]

Olsson, E. J., and Schubert, S. (2007), "Reliability Conducive Measures of Coherence", *Synthese* 157(3): 297-308.

Schlesinger, G. (1995), "Measuring Degrees of Confirmation", *Analysis* 55: 208-212.

Schubert, S. (2011), "Coherence and Reliability: The Case of Overlapping Testimonies", *Erkenntnis* 74(2): 263-275.

Shogenji, T. (1999). "Is Coherence Truth-conducive?", *Analysis* 59: 338-345.

PROPOSED READING Talbott, W. (2016), "Bayesian Epistemology", *Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/entries/epistemology-bayesian/ [A regularly updated overview of Bayesianism and its problems.]