Assessing Theories: The Coherentist Approach*

Peter Brössel
Institut für Philosophie II and
Center for Mind, Brain, and Cognitive Science
Ruhr-University Bochum

E-Mail: peter.broessel@rub.de

Contents

*A precursor of this paper appears as "Theory Assessment and Coherence" in the conference proceedings of a graduate conference at the Ruhr-University of Bochum in 2008. The conference proceedings appeared as A. Newen and T. Schlicht, special issue "In Honor of Rudolf Carnap: Central Issues in Epistemology" of *Abstracta*. I am especially grateful to Franz Huber for many valuable discussions on the topic of this paper and Bayesian epistemology in general. In addition, his exceedingly helpful comments on various drafts of this paper helped tremendously in clarifying the scope and the line of argumentation of the paper. I am also very thankful to Ralf Busse, Anna-Maria A. Eder, and two anonymous referees of this journal for their helpful remarks and suggestions.

Abstract

In this paper we show that the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003) satisfy the two most important adequacy requirements for the purpose of assessing theories. Following Hempel (1960), Levi (1967), and recently Huber (2008) we require, as minimal or necessary conditions, that adequate assessment functions favor true theories over false theories and true and informative theories over true but uninformative theories. We then demonstrate that the coherence measures of Olsson, Shogenji, and Fitelson satisfy these minimal conditions if we confront the hypotheses with a separating sequence of observational statements. In the concluding remarks we set out the philosophical relevance, and limitations, of the formal results. Inter alia, we discuss the problematic implications of our precondition that competing hypotheses must be confronted with a separating sequence of observational statements, which also leads us to discuss theory assessment in the context of scientific antirealism.

1 Introduction

Two of the most important questions in epistemology and the philosophy of science are: what is a good theory? and, when is a theory better than another theory, given some observational data? The coherentist's answer is the following: (*i*) a theory is a good theory given some observational data if and only if that theory coheres with the observational data, and (*ii*) a theory is better than another theory given some observational data if and only if the first theory coheres more with the observational data than the second. In this paper we ask whether this answer is adequate. More precisely, we investigate whether the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003) are adequate measures for the purpose of comparing and evaluating (i.e., assessing) theories.[1]

This application of coherence measures differs considerably from the intended application of coherence measures in the groundbreaking works of Bovens and Hartmann (2003) and Olsson (2005). The background is this: Many coherentists believe that the problem of the assessment of theories (or confirmation) and the problem of the justification of beliefs can be solved with the help of the concept of coherence. Traditionally both problems are closely related. They differ however in one important presupposition. Within the context of theory assessment or confirmation theory it is usually assumed that our information gathering processes are fully reliable and that we can rely fully on the data provided by these processes. This assumption is dropped when it comes to justification. In the context of justification

[1] The present paper's main purpose is to discuss coherence measures. Its main purpose is not to discuss confirmation measures, nor is it to discuss in detail which confirmation measures are adequate for the purpose of comparing and evaluating (i.e., assessing) theories. Although it would be interesting to compare the various coherence measures and the measures of confirmation proposed in the literature with respect to the question of how they fare as measures for theory assessment, we leave this for another occasion.

these information gathering processes might be only partially reliable. Bovens and Hartmann (2003) and Olsson (2005) study the coherence of different pieces of information provided by various independent and partially reliable witnesses, i.e., information gathering processes. Thus, what they ambitiously concentrate on is the problem of the justification of (all) our beliefs. However, the application considered here is the assessment of theories once the evidence (observational data that we can fully rely on) is given. Accordingly, it is assumed that our information gathering processes are fully reliable and that we can rely fully on the data provided by these processes. Glass (2007) considers a similar application of coherence measures. He "considers an application of work on probabilistic measures of coherence to inference to the best explanation" (Glass 2007, p. 275). In his paper "coherence is considered as a relation between an hypothesis and the evidence for it. Consequently, coherence relates the hypothesis to something which is already known to be true" (Glass 2007, p. 276). A further reason for the investigation of this kind of application is that there might be valuable lessons to learn for the construction of theories of justification when one considers the "simpler" but closely related problem of theory assessment first. A further question that must be left for another occasion is the following. Theories usually consist of different hypotheses that stand in various inter-theoretic relations to each other. Given that a theory is a conjunction of hypotheses $h_1, \ldots, h_n$, the question arises which effect the coherence of these hypotheses has upon the coherence of their conjunction (i.e. the theory) with the observational data. However, as already said answering this question is left for another occasion[2].

We begin, in section 2, by discussing what the necessary and sufficient conditions are that every assessment function must satisfy in order to be an adequate measure of the epistemic value of theories. Following Hempel 1960, Levi 1967, and Huber 2008 we require that adequate assessment functions favor true over false theories and that they favor true and informative theories over true but uninformative theories as minimal or necessary conditions. Furthermore, we discuss how adequate assessment functions should assess false theories. That good theories should be interpreted as true and informative theories is the content of a recent proposal in Huber's (2008) paper "Assessing Theories, Bayes Style."; parallels and discrepancies with Huber's theory will therefore be set out and discussed in this paper.

In section 3 we focus on the relation between coherence and theory assessment. As indicated above, we assume that we evaluate theories with the help of coherence measures by determining the degree of coherence of the theory with the available observational data. We show that the degree of coherence of a theory with the observational data depends on two indicators of the epistemic goodness of theories: the probability of the theory given the observational data, and its informativeness. These dependencies, worked out in section 3, help to prove in section 4 that the coherence measures of Olsson (2002), Shogenji (1999), and

[2]For the Shogenji measure of coherence this question is already answered in Brössel forthcoming.

Fitelson (2003) satisfy the minimal conditions for adequate measures of theory assessment, provided that we restrict ourselves to theories that do not contain theoretical vocabulary. Section 5 discusses theory assessment and the coherence measures in the context of scientific anti-realism. Section 6 summarizes and evaluates the achieved results.

2 What Is an Adequate Measure for Assessing Theories?

In this paper we investigate whether the coherence measures proposed by Olsson (2002), Shogenji (1999), and Fitelson (2003) are adequate measures for the purpose of assessing theories.[3] But what is an adequate measure for this purpose? The answer is as easy as can be, or so it seems. An adequate measure for assessing theories is a measure that leads us to theories that we would like to believe or accept. But which theories do we want to belief or accept? What is the goal of our scientific inquiry?

2.1 The Goal of Scientific Inquiry

According to Huber (2008) we want to believe or accept theories that are both: true and informative. A theory that we believe or accept should be true because we simply do not want to believe false theories. A theory that we believe or accept should be informative because from a theory we expect more than just its truth: a theory should also provides us with valuable information about how the world is—in other words, it possesses more than mere truth. The idea that we want to believe or accept true and informative theories was pursued by philosophers before Huber. The idea can be traced back to Popper (1935).[4] One of the clearest statements of this position can be found in an appendix to *The Logic of Scientific Discovery*, where Popper writes:

> *Science does not aim, primarily, at high probabilities. It aims at a high informative content, well backed by experience. But a hypothesis may be very probable simply because it tells us nothing, or very little.* A high degree of probability is therefore not an indication of "goodness" – it may be merely a symptom of low informative content. (Popper 1968, p. 399; emphasis in original)

A nice example that illustrates how frequently scientists prefer to accept or believe the more informative theory is provided by Salmon (2001).

[3]There are so many coherence measures in the literature that one cannot discuss all of them in one single paper, at least not in the formal detail I am aiming at. For example Douven and Meijs (2007) discuss several coherence measures that could not be considered here. The Bovens and Hartmann (2003) quasi-ordering of coherence is not taken under consideration because it is not a measure and it is defined in the rigid framework of testimonial systems.

[4]Popper 1954, Hempel 1960, Levi 1961 and 1967, and Hintikka and Pietarinen 1966 are the most important early papers that try to embody this idea in formal measures of confirmation.

> In general, the bolder a hypothesis is, the smaller its probability will be on any given body of evidence. If I predict (1) that it will rain in Pittsburg tomorrow, that is a more modest claim than (2) that at least 5 centimeters will fall, and that, in turn, is less bold than the statement (3) that between 5 and 10 centimeters will fall. Any set of conditions that satisfy (3) will necessarily satisfy (1) and (2), and any set of conditions that satisfy (2) will necessarily satisfy (1). Hence (1) is more probable than (2) and (2) is more probable than (3). However, (3), if true, is more informative than (2), and (2), if true, is more informative than (1). Scientists often choose bolder hypotheses because of their informational value, even if this means opting for less probable hypotheses. (Salmon 2001, p. 121)

These considerations show that theories which are true and informative are preferable to theories which are not true or not informative. However, this does not answer all our questions. Is every true theory preferable to every false theory (e.g., is a true but uninformative theory preferable to a false but informative theory)? How should we compare theories that have the same truth value (since it obviously cannot be the truth value which renders one theory preferable)? In some intuitive sense, true theories are better the more informative they are: is this also the case for false theories? The answers to these questions depend on what we mean by "information." So we have to answer the question: *When is a theory more informative than another theory?*

We must distinguish between a narrow and a wide usage of the term "information." According to the narrow usage of the term, information is always true; according to the wide usage, information is not always true. Only on the wide usage it is meaningful to speak of false information. In this paper we distinguish between "information" (in its wide usage), "true information," and "false information." Instead of using the term "information" we sometimes use the term "content." We never use the the term "information" in its narrow usage. The information or content of a theory is then the set of all sentences that are implied by the theory. The true information provided by a theory is the set of all true statements implied by that theory. The false information provided by a theory is the set of all statements implied by the theory that are not true. It should be noted that that if one true theory is logically stronger than another theory then the logically stronger theory provides more true information and is of greater content. However, it should also be noted that false logically stronger theories provide not only more false information they also provide more true information than logically weaker theories.[5]

[5] More formally: Let the true information provided by a theory T be the set $\mathfrak{True}\text{-}\mathfrak{Inf}(T) = \{A : A \textit{ is a true statement that is implied by } T\}$. Let the false information provided by a theory T is the set $\mathfrak{False}\text{-}\mathfrak{Inf}(T) = \{A : A \textit{ is a false statement that is implied by } T\}$. Suppose T_1 implies T_2. Hence, any conse-

In the following we briefly discuss several proposals as to which theories are to be preferred.

Proposal 1. *For possible world w, and theories T_1 and T_2, T_1 is to be preferred in w to T_2 if: (i) T_1 is true and T_2 is false in w, or (ii) T_1 and T_2 are true in w but T_1 is logically stronger (i.e., $T_1 \vdash T_2$ but $T_2 \nvdash T_1$), or (iii) T_1 and T_2 are false in w but T_1 is logically stronger (i.e., $T_1 \vdash T_2$ but $T_2 \nvdash T_1$). [see Huber 2008, p. 107]*

This is the proposal made in Huber 2008. According to (*i*) all true theories are to be preferred over all false theories. According to (*ii*) and (*iii*) if both theories have the same truth value the logically stronger theory is to be preferred. According to one reading of this proposal the logical strength or content of a theory is an epistemic value in its own right. The content of a theory contributes to the value of that theory independently of its truth value. However, logically stronger theories do not only possess more content but also provide more true information. Hence, another reading might be that the true information provided by a theory is of epistemic value and the false information provided by the theory plays no role, if both theories have the same truth value. Levi (1967, p. 77) suggests as much in reply to a proposal by Hempel (1960). A proposal that complies with Hempel's (1960) account is the following:

Proposal 2. *For possible world w, and theories T_1 and T_2, T_1 is to be preferred in w to T_2 if: (i) T_1 is true and T_2 is false in w, or (ii) T_1 and T_2 are true in w but T_1 is logically stronger than T_2, or (iii′) T_1 and T_2 are false in w but T_1 is logically weaker than T_2.*

This proposal differs from Proposal 1 with respect to comparing false theories only. If both theories are false this proposal focuses on the minimization of false information and, hence, the avoidance of error. It is important to note that these proposals do not dictate a complete ordering of how theories are to be preferred. Proposals 1 and 2 do not say anything about which theories are to be favored in the case that two theories have the same truth value but neither is logically implied by the other.

Both proposals share the requirements that (*i*) true theories are better than false theories and that (*ii*) logically stronger true theories are better than logically weaker theories. Requirement (*ii*) can be accepted without hesitation if one accepts (*i*). Presented with the choice between true theories we would always choose the logically stronger theory. (Claiming that we would choose the logically stronger theory of two true theories makes sense only if we presuppose a notion of theory choice along the lines of Hempel (1960): choosing a theory

---

quence of T_2 is also a consequence of T_1. This implies that if $A \in \mathfrak{True}\text{-}\mathfrak{Inf}(T_2)$ (i.e., if A is a true consequence of T_2) then $A \in \mathfrak{True}\text{-}\mathfrak{Inf}(T_1)$ (A is a true consequence of T_1) and if $A \in \mathfrak{False}\text{-}\mathfrak{Inf}(T_2)$ (i.e., if A is a false consequence of T_2) then $A \in \mathfrak{False}\text{-}\mathfrak{Inf}(T_1)$ (A is a false consequence of T_1). What is called true respectively false information here is called truth respectively falsity content by Popper (1968). Miller (1974) and Tichý (1974) provide a detailed discussion of the true and false content or information provided by theories.

T means adopting T as the logically strongest statement that one adds to the observational data E such that $T \wedge E$ is the logically strongest statement one believes or accepts. Thus, if one chooses theory $T_1 \wedge T_2$ one does not choose T_1, nor does one choose T_2.) Requirement (i) is often taken for granted in epistemology. If one wants to provide reasons, one may add that what theories are to be preferred depends on one's epistemic aims or purposes. If the question is what kind of theories we want to *believe* there is no plausible alternative to requirement (i). Believing a theory means that we accept that theory as true and if a theory is not true we certainly do not want to accept it as true. If, for example, the observational data implies that a theory is false, we would never believe that theory, no matter what other virtues it might have. However, some intuitions with respect to truth-likeness or verisimilitude and also empirical adequacy are in conflict with this requirement. In compliance with truth-likeness intuitions it may be the case that some false but very informative theories are better than some uninformative true theories; for example, Newtonian mechanics might be a better theory than a tautology, although we *know* that the first is false but the second is true. In the following we ignore theories of truth-likeness or verisimilitude since there is no commonly accepted proposal as to which theories are more truth-like than others. In compliance with intuitions with respect to empirical adequacy, requirement (i) is not plausible because it is not generally the case that false theories are worse than true theories. According to van Fraassen (1980) "[s]cience aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate"—not that it is true. If a false theory is empirically adequate it is not the case that all true theories are to be preferred to it. In particular, it may be the case that some informative empirically adequate but nevertheless false theories are better than some uninformative true theories. Discussion of the relationship between theory assessment, requirement (i), and empirical adequacy is taken up in more detail in section 5.

2.2 Adequate Measures for Assessing Theories

In the light of this discussion of the epistemic value of theories we can now investigate what we expect from a good measure for assessing theories. As already said, an adequate measure for assessing theories is a measure that leads us to good theories—respectively, is a means to the goal of our scientific inquiry. In other words, they are supposed to lead us to believe or accept those theories that we would prefer to believe or accept. More specifically, following Huber (2008) we require that an assessment function lead us "in the medium run" to true and informative theories, or, as he puts it, "to reveal the true assessment structure." Both proposals discussed above agree that true theories are preferable to false theories and that true but more informative theories are preferable to true but less informative theories, where a theory is taken to be more informative than another theory if it is logically stronger. Hence, we propose the following minimal conditions that every adequate assessment function must satisfy:

If a function $a : \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$ is an adequate assessment function in world w, then for all theories T_1, $T_2 \in \mathcal{L}$ and every data stream $e_1, \ldots, e_n, \ldots$ from $\mathcal{L}$ including all possible observational data of w, it holds that:

1. if T_1 is true in w and T_2 is false in w, then there exists a point n such that for all $m \geq n$:

$$a(T_1, E_m) > a(T_2, E_m)$$

2. if T_1 is true in w and T_2 is true in w and $T_1 \vdash T_2$ and $T_2 \nvdash T_2$, then there exists a point n such that for all $m \geq n$:

$$a(T_1, E_m) > a(T_2, E_m)$$

where $E_m = e_1 \wedge \ldots \wedge e_m$.

The first condition rests on the idea that we prefer to believe or accept true theories rather than false theories. Therefore an adequate assessment function—one which is supposed to quantify how good it is to believe or accept a theory in the light of the observational data—must favor a true theory over a false theory, at least after finitely many steps of observations, where one compares two theories, one of which is true and the other false. The second condition demands from an adequate assessment function that a more informative theory score higher than a less informative theory, given both theories are true. This second condition rests on the intuition that if we had to decide between two *true* theories that only provide true information, we would prefer to believe or accept the theory that provides more information. In the following section we consider on the one hand whether the coherence measures satisfy these minimal conditions, and on the other hand whether they agree or disagree with respect to the assessment of false theories.

3 Probability, Informativeness, and Coherence

In the previous section we saw what properties an adequate assessment function must exemplify. Accordingly, if we want to demonstrate that the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003) are such assessment functions we have to show that the degree of coherence of a theory with the available observational data depends somehow on the truth of the theory and the information it provides. In order to display such a dependence, the coherence of the theory with the observational data must somehow depend on an indicator of the theory's truth and an indicator of the theory's informativeness. The theory's probability with respect to the observational data, i.e., the theory's likeliness of truth, is the Bayesian's natural choice as an indicator of the theory's truth. As an indicator of the theory's informativeness or logical strength we introduce two measures of information.

3.1 Two Measures of Information

If we want to discuss in detail how we should measure the content of and the amount of true information provided by theories, fortunately, we can benefit from the work of other philosophers in this field. The first to develop a quantitative concept of information and to provide a theory thereof were Carnap and Bar-Hillel (see their technical report, Carnap and Bar-Hillel 1952). Hintikka (1999) puts the idea underlying their approach as follows:

> The basic idea of their [Carnap, Bar-Hillel] approach may be said to be one particular way of explicating the idea that information [in the wide usage of the word] equals elimination of uncertainty. In order to measure this uncertainty, a distinction is made between the different logical possibilities that we can express in a language. The more of them a statement s admits of, the more probable it is in some "purely logical" sense of probability. The more of them a statement s excludes, the less uncertainty it leaves, and the more informative will it therefore be. The probability $p(s)$ and information $inf(s)$ of a statement s are thus inversely related. (Hintikka 1999, p. 206)

The measure of content Carnap and Bar-Hillel suggested is the following.

Definition 3.1.

$$cont_p(T) = p(\neg T)$$

This measure of the content or information of a statement is highly plausible, since it satisfies a requirement which we might very well want to impose on the relation between content and probability. This requirement is the following: if $cont(T_1) - cont(T_2) = r$, then $p(T_2) - p(T_1) = a \times r + b$ for some $a, b \in \mathbb{R}$. This requires that if the content of a theory T_1 exceeds the content of a theory T_2 by r, then this should be reflected in the difference of the prior probabilities of the theories. A theory which has less content than another theory should not only be a priori more plausible, the difference in their a priori plausibilities should also be a function of their difference in content. More precisely, the difference in their probabilities should be an order-inverse and interval-preserving function of their difference in content. In the light of this requirement the above definition is the natural choice, if we additionally require that $cont : \mathcal{L} \rightarrow [0, 1]$ and $cont(A \vee \neg A) = 0$.

In the literature we can also find relativized measures of information or content. Since the observational data describes that part of the world that we are already familiar with via observations, we could assign degrees of informativeness to a theory relative to the observational data. Instead of asking how much a theory informs us about the world, we ask how much the theory informs us about the observational data. The following measure was proposed by Huber (2008) to measure exactly this.

Definition 3.2.

$$i_p(T, E) = p(\neg T|\neg E)$$

if $p(\neg E) > 0$

This information measure has an even longer history than the content measure. It was introduced by Hempel and Oppenheim six years before Carnap and Bar-Hillel wrote their technical report. In their paper *Studies in the Logic of Explanation* (1948) they presented it not as a measure of information but as a "measure of the systematic power" of a theory. It was meant to measure the explanatory and predictive power (in the sense of Hempel/Oppenheim) of a theory, given some data. In 1970 Risto Hilpinen suggested this measure for measuring how much the observational data informs us about a theory (accordingly, he used $p(\neg E|\neg T)$[6]). He called this the "normalized common content measure."[7] It is important to note that if we assume that the observational statement E is true, this measure is not only a measure of (relativized or normalized common) content but also a measure of true information delivered by T about E.

3.2 Probability, Informativeness, and Coherence

In the following we investigate the relation between the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003), and the probability and content or informativeness of a theory in the light of the observational data.

3.2.1 Probability, Informativeness, and Olsson's Measure of Coherence

Olsson's (2002, p. 250) measure of coherence is defined as follows

Definition 3.3.

$$\mathcal{C}_{O,p}(A_1, \ldots, A_n) = \frac{p(\bigwedge_{1\leq i\leq n} A_i)}{p(\bigvee_{1\leq i\leq n} A_i)}$$

[6] As already explained, $i_p(T, E)$, respectively $p(\neg T|\neg E)$, quantifies how much the theory informs us about the observational data. Since Hilpinen (1970) is interested in quantifying the degree of information provided by the observational data about the theory, he uses $i_p(E, T)$, respectively $p(\neg E|\neg T)$.

[7] Hilpinen calls it the normalized common content measure because of its relation to the content measures $cont_p$. Formally the relation is the following: $i_p(T, E) = \frac{cont_p(T\vee E)}{cont_p(E)}$. Hilpinen interprets $cont_p(T \vee E)$ as a measure of the common content of T and E. Hence, $i_p(T, E)$ relates the common content of T and E to the overall content of E and quantifies what proportion of the content of E is also content of T. See Hilpinen (1970) for a detailed axiomatic motivation of this information measure.

if $p(\bigvee_{1\leq i\leq n} A_i) > 0$ and 0 otherwise.

According to Olsson (2002) this measure of coherence measures the degree of agreement of the statements $A_1, \ldots, A_n$. However, a far more obvious interpretation of Olsson's coherence measure is that it measures how much the propositions expressed by $A_1, \ldots, A_n$ hang together. Here we understand "hang together" in the following way: propositions hang together if they are either true together or false together. In other words, propositions hang together if, if at least one of the propositions is true then all of them are true. Olsson's coherence measure fits this intuition perfectly. The degree of coherence of statements $A_1, \ldots, A_n$ equals the conditional probability of all of them being true, given that at least one of them is true, i.e., their disjunction is true.[8]

Now let us take a closer look at the connection between the degree of coherence of a theory and the observational data, probability, and the above-mentioned measures of content and information. From the definition of Olsson's coherence measure and the fact that $p(T \vee E) = 1 - p(\neg T \wedge \neg E)$ it is easy to obtain the following result:

Theorem 3.1. If $p(T_1|E) = p(T_2|E) > 0$, then:[9]

$$\mathcal{C}_{O,p}(T_1, E) > \mathcal{C}_{O,p}(T_2, E) \textit{ iff } i_p(T_1, E) > i_p(T_2, E).$$

And we also can prove the following:

Theorem 3.2. If $i_p(T_1, E) = i_p(T_2, E)$, then:

$$\mathcal{C}_{O,p}(T_1, E) > \mathcal{C}_{O,p}(T_2, E) \text{ iff } p(T_1|E) > p(T_2|E).$$

Theorem 3.1 shows that if the conditional probabilities of two theories T_1 and T_2 equal each other given the observational data E, then, according to Olsson's definition the degree of coherence, of T_1 and E is higher than the degree of coherence of T_2 and E iff T_1 informs more (in the sense of $i_p(\cdot, \cdot)$) about the observational data E than T_2. This shows that the degree of coherence between a theory and the observational data indeed depends on the informativeness of the theory. Theorem 3.2 shows that if we fix the amount of information provided by the

[8] This is obvious since $\models (A_1 \wedge \ldots \wedge A_n) \leftrightarrow (A_1 \wedge \ldots \wedge A_n) \wedge (A_1 \vee \ldots \vee A_n)$. This implies that $\frac{p(A_1 \wedge \ldots \wedge A_n)}{p(A_1 \vee \ldots \vee A_n)} = p(A_1 \wedge \ldots \wedge A_n | A_1 \vee \ldots \vee A_n)$. It is remarkable and astonishing that our understanding of "hanging together" is the same as that of Shogenji (1999). He writes: "The crudest way of unpacking the idea that coherent beliefs 'hang together' is that they are either true together or false together. However, coherence comes in degrees; in other words we want to say that the *more* coherent beliefs are, the *more* likely they are true together" (Shogenji 1999, p. 338). But Shogenji does not agree with our formal interpretation of this intuitive notion of "hanging together," since he concludes that "[t]he more coherent two beliefs are, the stronger is the positive impact of the truth of one on the truth of the other" (Shogenji 1999, p. 338).

[9] Please note, $p(T|E) > 0$ implies not only that $p(T|E)$ is defined, but also that $p(E) > 0$.

theories relative to the observational data, then that theory which coheres more with the observational data is more probable in the light of the observational data.

Both results fit our intuitions perfectly. If we had to decide between equally probable theories, given the data E, we would choose the more informative one. We would choose the more informative one because we would get more information with the same risk of accepting a false theory. And if we had to choose between equally informative theories we would choose the more probable one, because this one is more likely to be true.

The following theorem shows a more general result concerning the relation between the presented information measures $i_p(\cdot,\cdot)$ and the conditional probability on the one hand, and the degree of coherence of a theory and the observational data on the other.

Theorem 3.3. $\forall p\ \forall \epsilon > 0\ \exists \delta_\epsilon > 0$: if $[p(T_1|E) > 0\ \&\ i_p(T_1, E) \geq i_p(T_2, E) + \epsilon\ \&\ p(T_1|E) \geq p(T_2|E) - \delta_\epsilon]$, then $\mathcal{C}_{O,p}(T_1, E) > \mathcal{C}_{O,p}(T_2, E)$.

This theorem shows that Olsson's coherence measure weighs the the two epistemic virtues of a theory—i.e., its likeliness of truth and its informativeness. A theory T_1 can cohere more with the observational data than another theory T_2 even if the conditional probability of the latter is higher than the conditional probability of the former. This happens if the amount of information about the observational data E provided by T_1 is sufficiently higher than the amount of information provided by T_2. This shows that a more informative theory can display a higher degree of coherence with the observational data than a less informative theory. It suffices that the difference between the conditional probabilities of both theories is small enough relative to the difference in their informational value. It should be recognized that this theorem instantiates a further condition which Huber imposes on plausibility-informativeness assessment functions, namely that "any surplus in informativeness succeeds, if the shortfall in plausibility is small enough" (Huber 2008, p. 94). Huber dubs this requirement *continuity.*

We can conclude that the degree of coherence between a theory and the observational data depends on both indicators of the epistemic virtues of theories, at least with respect to Olsson's measure of coherence: the theory's probability and its informativeness. With this result we have a first hint that Olsson's measure of coherence indeed satisfies the necessary requirements for adequate assessment functions.

3.2.2 Probability, Informativeness, and Shogenji's Measure of Coherence

For Shogenji's measure of coherence we can prove very similar theorems as for Olsson's measure. Shogenji (1999) introduces his measure of coherence in the following way.

Definition 3.4.

$$\mathcal{C}_{S,p}(A_1, \ldots, A_n) = \frac{p(\bigwedge_{1\leq i\leq n} A_i)}{\prod_{1\leq i\leq n}(p(A_i))}$$

if $p(A_i) > 0$ for all $i : 1 \leq i \leq n$, and 0 otherwise.

Shogenji (1999), unlike Olsson (2002), also defines the term "the statements $A_1, \ldots A_n$ are coherent." He defines this as follows:

Definition 3.5. The statements $A_1, \ldots A_n$ are coherent iff

$$\mathcal{C}_{S,p}(A_1, \ldots, A_n) > 1$$

The underlying intuition of definitions 3.4 and 3.5 is that the coherence of statements depends on how much the statements mutually support each other. They are coherent if there is at least some positive probabilistic dependency between them. And they are more coherent than other statements if the positive probabilistic dependencies between them surpass the positive probabilistic dependencies of the others.

As already said, for Shogenji's measure of coherence we can prove similar theorems as for Olsson's measure.

Theorem 3.4. If $p(T_1|E) = p(T_2|E) > 0$, then:

$$\mathcal{C}_{S,p}(T_1, E) > \mathcal{C}_{S,p}(T_2, E) \text{ iff } cont_p(T_1) > cont_p(T_2).$$

And we also can prove the following result:

Theorem 3.5. If $cont_p(T_1) = cont_p(T_2) > 0$, then:

$$\mathcal{C}_{S,p}(T_1, E) > \mathcal{C}_{S,p}(T_2, E) \text{ iff } p(T_1|E) > p(T_2|E).$$

The comments on these theorems are the same as for the theorems 3.1 and 3.2. Both theorems show that, according to Shogenji's coherence measure, the degree of coherence of a theory and the observational data depends on the theory's probability and its content as specified by the content measure $cont_p(\cdot)$; thus, again, both results fit our intuitions perfectly. If we had to decide between equally probable theories, given the data E, we would choose the theory with more content. We would choose the theory with more content because we could believe more with the same risk of accepting a false theory. And if we had to choose between theories with same amount of content, we would choose the more probable one, because this one is more likely to be true.

Again, we can prove a more general result which indicates that Shogenji's coherence measure weighs the probability and the content of a theory.

Theorem 3.6. $\forall p \ \forall \epsilon > 0 \ \exists \delta_\epsilon > 0$: if $cont_p(T_1) \geq cont_p(T_2) + \epsilon$ & $0 \neq p(T_1|E) \geq p(T_2|E) - \delta_\epsilon]$, then $\mathcal{C}_{S,p}(T_1, E) > \mathcal{C}_{S,p}(T_2, E)$.

We can conclude that according to Shogenji's measure of coherence the degree of coherence between a theory and the observational data depends on the theory's probability and its content. This proves that there is a dependency between the degree of coherence of a theory and the observational data and the theory's probability given the observational data, and the theory's content. Additionally, the above result shows that Shogenji's coherence measure fulfills Huber's continuity requirement with respect to the theory's content (in the sense of $cont_p(\cdot)$).

A remarkable difference between the theorems 3.1, 3.2, and 3.3 on the one hand and theorems 3.4., 3.5, and 3.6 on the other is the following: the degree of coherence according to to theorems 3.1, 3.2, and 3.3 concerning Olsson's measure depends on the informativeness of a theory as specified in the information measure $i_p(\cdot,\cdot)$ (which measures the informativeness of a theory about some observational data), whereas the degree of coherence according to theorems 3.4 , 3.5, and 3.6 concerning Shogenji's measure depends on the content of a theory as specified by the measure $cont_p(\cdot)$. However, it is important to note that these theorems do not already establish conclusively that that Olsson's measure depends on the informativeness of a theory as specified in the information measure $i_p(\cdot,\cdot)$ (which measures the informativeness of a theory about some observational data), whereas the degree of coherence according Shogenji's measure depends on the content of a theory as specified by the measure $cont_p(\cdot)$. In particular, one can show that if $p(T_1|E) = p(T_2|E) > 0$, then $cont_p(T_1) > cont_p(T_2)$ iff $i_p(T_1, E) > i_p(T_2, E)$. However, it would be too hasty to conclude from this that there is only seemingly a difference between Olsson's and Shogenji's measure. In particular, note the following theorem:

Theorem 3.7. There is some probability function p such that $p(T_1|E_1) = p(T_2|E_2) > 0$ and

$$\mathcal{C}_{O,p}(T_1, E_1) > \mathcal{C}_{O,p}(T_2, E_2) (\text{i.e. if } i_p(T_1, E_1) > i_p(T_2, E_2))$$

but

$$\mathcal{C}_{S,p}(T_1, E_1) < \mathcal{C}_{S,p}(T_2, E_2) (\text{i.e. if } cont_p(T_1) > cont_p(T_2))$$

Similarly theorems can be shown if we want to hold fixed $cont_p(T_1) = cont_p(T_2)$ or $i_p(T_1, E_1) = i_p(T_2, E_2)$ and vary $p(T_1|E)$ and $p(T_2|E)$. This demonstrates that there is indeed a fundamental difference between the two measures. Olsson's measure depends on the informativeness of a theory as specified in the information measure $i_p(\cdot,\cdot)$, whereas the degree of coherence according to Shogenji's measure depends on the content of a theory as specified by the measure $cont_p(\cdot)$.

3.2.3 Probability, Informativeness, and Fitelson's Measure of Coherence

Fitelson (2003) proposes the following coherence measure.

Definition 3.6.

$$\mathcal{C}_{F,p}(A_1, \ldots, A_n) = \frac{1}{|\mathcal{R}|} \sum_{\langle A_i, S_j \rangle \in \mathcal{R}} \mathcal{F}(A_i, \bigwedge_{A_m \in S_j} A_m)$$

where $\mathcal{R} = \{\langle A_i, S_j \rangle : A_i \in \{A_1, \ldots A_n\} \text{ and } S_j \subseteq \{A_1, \ldots A_n\} \text{ and } A_i \notin S_j \text{ and } S_j \neq \varnothing\}$.

In his definition of a coherence measure Fitelson uses a variation of Kemeny-Oppenheim's *measure of factual support* $\mathcal{F}$, which they introduced in their 1952 joint paper as a measure of mutual support.[10]

Definition 3.7.

$$\mathcal{F}(A_1, A_2) = \frac{p(A_2|A_1) - p(A_2|\neg A_1)}{p(A_2|A_1) + p(A_2|\neg A_1)}$$

if $0 < p(A_1) < 1$ and $p(A_2) > 0$, otherwise if $p(A_1) = 0$ or $p(A_2) = 0$, then $\mathcal{F}(A_1, A_2) = -1$ and if $p(A_1) = 1$ and $p(A_2) > 0$, then $\mathcal{F}(A_1, A_2) = 1$.

Again we can prove similar theorems for Fitelson's measure as for Olsson's and Shogenji's. First note that:

Theorem 3.8. If $p(T_1|E) = p(T_2|E) > 0$, then:

$$\mathcal{C}_{F,p}(T_1, E) > \mathcal{C}_{F,p}(T_2, E) \text{ iff } cont_p(T_1) > cont_p(T_2).$$

We conclude that according to Fitelson's coherence measure it is also true that the degree of coherence of a theory and the observational data depends on the content of the theory. For two theories which are equally probable in the light of the observational data, it holds that the theory with more content coheres more with the observational data than the theory with less content.

Theorem 3.9. If $cont_p(T_1) = cont_p(T_2) < 1$, then:

$$\mathcal{C}_{F,p}(T_1, E) > \mathcal{C}_{F,p}(T_2, E) \text{ iff } p(T_1|E) > p(T_2|E).$$

This theorem shows that if we consider two theories with the same amount of content, then the theory that coheres more with the observational data is more probable in the light of the observational data.

Theorem 3.10. $\forall p\ \forall \epsilon > 0\ \exists \delta_\epsilon > 0:$ if $[p(T_1 \wedge E) > 0\ \&\ cont_p(T_1) \geq cont_p(T_2) + \epsilon\ \&\ p(T_1|E) \geq p(T_2|E) - \delta_\epsilon]$, then $\mathcal{C}_{F,p}(T_1, E) > \mathcal{C}_{F,p}(T_2, E)$.

Again we get that the degree of coherence of a theory and the observational data depends on both the conditional probability of the theory and the content of the theory. This is shown by the fact that Fitelson's measure weighs both aspects to determine the degree of coherence of theory and observational data. The last theorem also shows that the Fitelson coherence measure fulfills Huber's continuity requirement with respect to the content measure $cont_p(\cdot)$.

[10] For a detailed argument in support of the Kemeny-Oppenheim measure of factual support, see Fitelson 2001, esp. sect. 3.2.3.

4 Theory Assessment and Coherence

4.1 Theory Assessment, Coherence, and True Theories

In this section we show that the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003) satisfy the necessary requirements on adequate assessment functions. Therefore we have to show that they fulfill the requirements on adequate assessment functions we laid down in section 2.2. We required that they favor true theories over false theories and that they favor logically stronger true theories over logically weaker true theories, since the former are more informative than the latter. In the last section we already saw a hint that all three coherence measures have this property. Theorems 3.3, 3.6, and 3.10 showed that a more informative theory, or a theory with greater content, can cohere more with the observational data if the difference in their conditional probabilities is small enough. Now the basic idea for the proof that the coherence measures are good assessment functions is the following: First, the content and the information provided by a logically stronger, true theory is always higher than the content and the information provided by a logically weaker, true theory. Second, by the Gaifman-Snir Theorem (1982) we know that the conditional probability of a true theory tends towards its truth-value if confronted with a separating sequence of statements. (To be more exact, the sequence must separate the set of possibilities $Mod_{\mathcal{L}}$. A sequence of statements $e_1, \ldots, e_n, \ldots$ separates the set of possibilities $Mod_{\mathcal{L}}$ if and only if for every pair of worlds w_i and $w_j \in Mod_{\mathcal{L}}$ (with $w_i \neq w_j$) there is one statement in the sequence such that it is true in one of the possible worlds and false in the other.)

Suppose we confront two true theories with such a sequence of observational data. Then the conditional probabilities of both theories tend to 1 since those theories are true. As a consequence the difference between their conditional probabilities becomes smaller and smaller, since their conditional probabilities tend to 1. This opens up the possibility that the more informative theory, or the theory with greater content, will cohere more with the separating observational data than the less informative theory, or the theory with less content.

The theorem which shows that the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003) are indeed good assessment functions is the following:

Theorem 4.1. Let $w \in Mod_{\mathcal{L}}$ be a possible world and let $e_1, \ldots, e_n, \ldots$ be a sequence of statements of $\mathcal{L}$ which separates $Mod_{\mathcal{L}}$, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let p be a strict (or regular) probability function on $\mathcal{L}$. Let p^* be the unique probability function on the smallest σ-field $\mathcal{A}$ containing the field $\{Mod(A) : A \in \mathcal{L}\}$ satisfying $p^*(Mod(A)) = p(A)$ for all $A \in \mathcal{L}$, where $Mod(A) = \{w \in Mod_{\mathcal{L}} : w \vDash A\}$ and $Mod_{\mathcal{L}}$ is the set of all maximal-consistent sets of statements of $\mathcal{L}$ including instances.

Then there is an $X \subseteq Mod_{\mathcal{L}}$ with $p^*(X) = 1$ such that the following holds for every $w \in X$

and all theories T_1 and T_2 of $\mathcal{L}$:

If $\mathcal{C} = \mathcal{C}_{O,p}$ or $\mathcal{C} = \mathcal{C}_{S,p}$ or $\mathcal{C} = \mathcal{C}_{F,p}$, then

1. if $w \models T_1$ and $w \models \neg T_2$, then:

   $$\exists n \forall m \geq n : [\mathcal{C}(T_1, E^w_m) > \mathcal{C}(T_2, E^w_m)].$$

2. If $w \models T_1 \wedge T_2$ and $T_1 \vdash T_2$ but $T_2 \nvdash T_1$, then:

   $$\exists n \forall m \geq n : [\mathcal{C}(T_1, E^w_m) > \mathcal{C}(T_2, E^w_m)]$$

   where $E^w_m = \bigwedge_{0 \leq i \leq m} e^w_i$.

This theorem shows that if we compare two theories, one of them true and one of them false, then the true theory coheres more with the observational data than the false theory (after finitely many steps of observation in a sequence of separating observational statements and for every observation thereafter).[11] And it also shows that if we compare two theories, both of them true but one of them logically stronger, then the logically stronger theory—i.e., the more informative theory—coheres more with the observational data (after finitely many steps of observation and for every observation thereafter). Here the fact that we compare logically stronger, true theories with logically weaker, true theories is the reason why we can show that all three coherence measures satisfy the minimal conditions for being good theory assessment functions. Olsson's (2002) measure of coherence weighs between the plausibility and the informativeness of the theory. Shogenji's (1999) measure of coherence and Fitelson's (2003) measure of coherence weigh between the theory's plausibility and its content. However, in the case of one logical stronger, true theory and one logically weaker, true theory the more informative theory is necessarily also the one with more content and *vice versa*. Hence, the measures agree in the comparative evaluation of the logically stronger, true theory and the

[11]For the coherence measures of Shogenji (1999) and Fitelson the proofs in the appendix actually show the following stronger result: all true theories cohere with the observational data after finitely many steps of observations and for every observation thereafter; false theories don't . No such theorem is provable for Olsson's (2002) measures of coherence since coherence is not defined for this measure. Nevertheless Olsson's coherence measure satisfies the minimal conditions on good theory assessment functions put forward in section2.2. The reason for this is that the minimal conditions on good theory assessment functions only require that if one consider two theories, on of which is true and the other false, then the true theory cohere more strongly with the evidence than the false theory. This comparative requirement can be satisfied even if their is no definition of the qualitative notion of coherence available.

logically weaker, true theory. Olsson's measure favors logically stronger, true theories because they are more informative in the sense of measure $i_p(\cdot|\cdot)$. Shogenji's coherence measure and Fitelson's coherence measure favor logically stronger, true theories because they have more content in the sense of measure $cont_p(\cdot)$. This holds for all theories even if the theory can never be entailed or refuted by the observational data after finitely many steps of observation in a sequence of separating observational statements.

4.2 Theory Assessment, Coherence, and False Theories

Now let us turn to the question of how the different coherence measures assess false theories. Proposal 1 states that if two theories have the same truth value the logically stronger theory is to be preferred. According to Proposal 2 false logically stronger theories should be rated lower than false logically weaker theories. None of the introduced coherence measures complies with these proposals. Instead they rate all false theories equal *in the limit*. The degree of coherence of false theories converges to the respective minimal degree of coherence of that measure. This is shown in the following theorem:

Theorem 4.2. Let $e_1, \ldots, e_n, \ldots$ be a sequence of statements of $\mathcal{L}$ which separates $Mod_{\mathcal{L}}$, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let p be a strict (or regular) probability function on $\mathcal{L}$. Let p^* be the unique probability function on the smallest σ-field $\mathcal{A}$ containing the field $\{Mod(A) : A \in \mathcal{L}\}$ satisfying $p^*(Mod(A)) = p(A)$ for all $A \in \mathcal{L}$, where $Mod(A) = \{w \in Mod_{\mathcal{L}} : w \vDash A\}$ and $Mod_{\mathcal{L}}$ is the set of all maximal-consistent sets of statements of $\mathcal{L}$ including instances.
Then there is an $X \subseteq Mod_{\mathcal{L}}$ with $p^*(X) = 1$ such that the following holds for every $w \in X$ and all theories T_1 of $\mathcal{L}$:

If $w \vDash \neg T_1$, then:

$$lim_{n\to\infty}[\mathcal{C}_{O,p}(T_1, E_m^w)] = 0$$

$$lim_{n\to\infty}[\mathcal{C}_{S,p}(T_1, E_m^w)] = 0$$

$$lim_{n\to\infty}[\mathcal{C}_{F,p}(T_1, E_m^w)] = -1$$

where $E_m^w = \bigwedge_{0\leq i\leq m} e_i^w$.

Thus, false theories are rated equal *in the limit* by all three coherence measures. In particular, logically false theories are assigned immediately the minimal degree of coherence with the observational data according to all three coherence measures. Falsifiable false theories are falsified after finitely many steps of observation in a sequence of separating observational statements and are assigned the minimal degree of coherence for every observation thereafter.

The degree of coherence with the observational data of false theories that cannot be falsified converges to the minimal degree but never actually reaches it. The existence of false theories that cannot be falsified after finitely many steps of observation is the reason why not all false theories are rated equal after finitely many steps of observation.

4.3 Theory Assessment and Coherence: Constraints

The achieved results hold true if two conditions are satisfied. First, the observational data must be of such a kind that they separate the set $Mod_{\mathcal{L}}$. Second, they hold true only in every world w of some subset $X \subseteq Mod_{\mathcal{L}}$ with $p^*(X) = 1$. That the latter condition must be satisfied is a problem of Bayesian updating of probability functions in general. (See, for example, chapter 6 of Earman (1992) for further discussion of this problem.)

The condition that the observational data must separate the set $Mod_{\mathcal{L}}$ is problematic, too. This means that theorem 4.1 does not speak about theories which are formulated in theoretical vocabulary, i.e., vocabulary which includes non-observational terms; but such theories are common in everyday scientific practice.

However, there are some philosophers who do not think this is problematic, since they claim that an anti-realist position with respect to scientific theories is preferable.

5 Theory Assessment, Coherence, and Scientific Anti-Realism

According to van Fraassen (1980, p. 9) Scientific Realism is "the position that scientific theory construction aims to give us a literally true story of what the world is like, and that acceptance of a scientific theory involves the belief that it is true." Advocates of an anti-realist point of view deny this. Van Fraassen mentions two alternative sorts of anti-realist positions.

The first sort of anti-realist position is that scientific theories are not literally true or false, but are true or false if properly construed. According to this position, theoretical terms within scientific theories do not refer in the way other terms do, and statements containing them are not literally true or false. A sentence, properly construed, which can be true or false must therefore be a sentence which does not include any theoretical terms.

The second sort of anti-realist position is that scientific theories are literally true or false, but the acceptance of a theory does not involve the belief that it is true. Instead, this anti-realist says, we accept theories because the theory possesses other virtues, such as empirical adequacy. Anti-realists of both sorts should not have problems with the condition that the observations must separate $Mod_{\mathcal{L}}$.

Anti-realists of the first sort claim that theories which are literally true or false do not contain theoretical vocabulary. Since they also claim that scientific theories, properly construed, are true or false, they must hold that scientific theories, properly construed, do not contain theoretical terms. This means that theories, properly construed, can be formulated in a language $\mathcal{L}_{obs}$ that does not include non-observational terms. Accordingly, there is a sequence of observational statements that separates the set of possible worlds or models $Mod_{\mathcal{L}_{obs}}$ of the language $\mathcal{L}_{obs}$. Therefore anti-realists of the first sort can use coherence measures to assess theories. They can use the coherence measures to assess theories because theorem 4.1 shows that if we compare two *properly construed* theories, one of them true and one of them false, then the true theory coheres more with the observational data than the false theory (after finitely many steps of observation in a sequence of separating observational statements and for every observation thereafter). And it also shows that if we compare two *properly construed* theories, both of them true but one of them logically stronger, then the logically stronger theory—i.e., the more informative theory—coheres more with the observational data (after finitely many steps of observation in a sequence of separating observational statements and for every observation thereafter). So an anti-realist of the first sort need not have any problem with the condition that the observational data must separate the set $Mod_{\mathcal{L}}$.

An anti-realist of the second sort will not have any problem with this condition either. Since according to her position accepting a theory does not involve believing that it is true, only that it is empirically adequate.

Suppose all theories under consideration are formulated in the language $\mathcal{L}$ which contains observational as well as theoretical terms. Now let $\mathcal{L}_{obs}$ be a proper subset of the language $\mathcal{L}$ which contains all and only those statements which are formulated in observational terms. Suppose theory T is formulated in language $\mathcal{L}$ and contains theoretical terms, and suppose theory T^{obs} is the logically strongest theory that is formulated in language $\mathcal{L}_{obs}$ and which is implied by T. In addition assume that the quantifiers in T^{obs} range exclusively over the set of observable entities. The reason for this later assumption is that if the quantifies are unrestricted "we shall be able to state in the observational language that there are unobservable entities" (van Fraassen 1980, p. 54). According to van Fraassen a theory is "empirically adequate exactly if what it says about the observable things and events in the world is true—exactly if it 'saves the phenomena' "(van Fraassen 1980, p. 12). Accordingly, the following should hold true: T is empirically adequate iff T^{obs} is true. In addition we want to introduce the comparative concept "is more empirically adequate than," the following minimal requirement seems reasonable: If T_1 and T_2 are both empirically adequate, and T_1^{obs} is logically stronger than T_2^{obs}, then T_1 is more empirically adequate than T_2.[12]

[12]With the help of the content measure *cont* and theory T^{obs} one can introduce a measure of the empirical content of a theory T instead of just the content of a theory. In particular, $cont_p(T^{obs})$ can be said to measure the amount of *empirical* content of T.

Along with introducing the comparative concept "is more empirically adequate than," an anti-realist of the second sort would also reformulate the above-mentioned minimal conditions that every adequate assessment function must satisfy:

If a function $a : \mathcal{L} \times \mathcal{L} \to \mathbb{R}$ is an adequate assessment function in world w, then for all theories T_1, $T_2 \in \mathcal{L}$ and every data stream $e_1, \ldots, e_n, \ldots$ from $\mathcal{L}_{obs}$ including all possible observational data of w, it holds that:

1. if T_1 is empirically adequate in w and T_2 is empirically inadequate in w, then there exists a point n such that for all $m \geq n$:

$$a(T_1, E_m) > a(T_2, E_m)$$

2. if T_1 is empirically adequate in w and T_2 is empirically adequate in w and $T_1^{obs} \vdash T_2^{obs}$ and $T_2^{obs} \nvdash T_2^{obs}$, then there exists a point n such that for all $m \geq n$:

$$a(T_1, E_m) > a(T_2, E_m)$$

where $E_m = e_1 \wedge \ldots \wedge e_m$.

There are two ways anti-realists of the second sort can employ the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003) for the assessment of theories.

First, anti-realists of the second sort can determine the degree of coherence $Coh(T, E)$ of the theory T and the observational data to assess theory T by defining $a(T, E) := Coh(T, E)$. In that case one *cannot* show that the coherence measures favor empirically adequate theories over empirically inadequate theories and one also *cannot* show that the coherence measures favor true theories over false theories. The reason is that if the language $\mathcal{L}$ in which T is formulated contains theoretical terms, the observational statements cannot separate the set of possibilities $Mod_{\mathcal{L}}$ and hence we cannot use theorem 4.1 to assess theories.[13] That coherence measures cannot be used to assess theories directly is a surprising result since intuitively, and historically, coherence theories are closely tied to anti-realist positions in the philosophy of science and epistemology.

[13]It is important to note that this utilization of coherence measures to assess theories is inadequate from the perspective of an anti-realist of the second sort *even if* the coherence measures would favor true theories over false theories. The reason is that anti-realists of the second sort do not prefer true theories over false theories in general. In particular, if both theories are empirically adequate an anti-realist of the second sort would not necessarily favor the true theory over the false theory. In addition, anti-realists of the second sort would reject the consequence (Theorem 4.2) that all false theories should be treated on a par in the long run. After all, anti-realists of the second sort subscribe to the point of view that false but empirically adequate theories are to be preferred to false and empirically inadequate theories.

Second, however, anti-realists of the second sort can use the coherence measures to assess theories indirectly. In particular, they can determine the degree of coherence $Coh(T^{obs}, E)$ of theory T^{obs} and observational data E to assess theory T by defining $a(T, E) := Coh(T^{obs}, E)$, where T^{obs} is again the logically strongest theory that is formulated in language $\mathcal{L}_{obs}$ and which is implied by T. In that case Theorem 4.1 shows that if T_1 is empirically adequate and T_2 is not, i.e., if T_1^{obs} is true and T_2^{obs} is false, then after finitely many steps of observations T_1^{obs} coheres more with the observational data than T_2^{obs}. By theorem 4.1 we then know that if T_1 and T_2 are both empirically adequate—i.e., if both T_1^{obs} and T_2^{obs} are true—then after finitely many steps of observations T_1^{obs} coheres more with the observational data than T_2^{obs} if T_1^{obs} is logically stronger than T_2^{obs}. By evaluating theories which do not contain theoretical terms (e.g., T_1^{obs} and T_2^{obs}) we can therefore determine which theories containing theoretical terms (e.g., T_1^{obs} and T_2^{obs}) are empirically adequate, and which are more empirically adequate than others.

This shows that an anti-realist of the second sort can use the coherence measures of Olsson (2002), Shogenji (1999), and Fitelson (2003) to assess theories without worrying about the precondition that the observational data must separate the set of possibilities $Mod_{\mathcal{L}}$. However, anti-realists of the second sort have to assess a theory T indirectly by determining the degree of coherence between T^{obs} and the observational data E, where T^{obs} is the logically strongest theory implied by T not containing theoretical vocabulary.[14] We conclude: at least from the perspective of anti-realists of this second sort, the coherence measures satisfy the minimal conditions for adequate assessment functions. If one compares two theories, one of which is empirically adequate and the other not, they favor the empirically adequate theory over the empirically inadequate theory and they favor the theory with more empirical content among two empirically adequate theories after finitely many steps of observation and for every observation thereafter.

6 Summary

We have shown in section 4.1 that the coherence measures of Olsson, Shogenji, and Fitelson satisfy the two most important requirements for adequate theory assessment functions laid down in section 2.2, provided that the possible observational data separates the set of all

[14]Van Fraassen (1980) propagates this second sort of anti-realist position most prominently. One might object that van Fraassen advocates a view on theories which stands in sharp contrast to the one presumed here. However, as noted by Hawthorne (2011, p. 333), van Fraassen's *semantic view of theories* need not be in opposition to the assumption that scientific theories are expressible in a sentence or statement of some language. "Presumably, if scientists can express a theory well enough to agree about what it says about the world (or at least about its testable empirical content), it must be expressible in some bit of language." Such a theory "should be subject to empirical evaluation [...and] a theory of confirmation [or theory assessment] should apply to them."

possibilities $Mod_{\mathcal{L}}$. (i) If we compare two theories (formulated within the vocabulary of $\mathcal{L}$), one of them true and one of them false, then the true theory coheres more with the observational data than the false theory (after finitely many steps of observation in a sequence of separating observational statements and for every observation thereafter) according to all three coherence measures. (ii) If we compare two theories, both of them true but one of them logically stronger, then the logically stronger theory—i.e., the more informative theory—coheres more with the observational data (after finitely many steps of observation in a sequence of separating observational statements and for every observation thereafter) according to all three coherence measures. However, this result comes with a caveat. As already noted in section 4.3, this result does not speak about theories which are formulated using theoretical vocabulary. If we assume that the sequence of possible observational data separates the set of possibilities $Mod_{\mathcal{L}}$, and if the theories must be formulated exclusively with the help of the vocabulary of $\mathcal{L}$, then the theories cannot be formulated using theoretical vocabulary. Accordingly, if you are a scientific realist and aim at believing true theories, the coherence measures do not necessarily help you reach that goal, in particular not if you concentrate on theories that are formulated in theoretical vocabulary.

If you are scientific anti-realist you do not have to worry that the achieved results do not speak about theories that contain theoretical vocabulary. If you belong to the first type of anti-realists, you believe that theories which are literally true or false do not contain theoretical terms, and that theories properly construed are true or false; hence, you also hold that scientific theories, properly construed, do not contain theoretical terms. Accordingly you can apply the coherence measures to theories properly construed and they will take you to true theories. If you belong to the second type of anti-realists you do not aim at believing true theories, rather you aim at believing theories which are empirically adequate, i.e., theories whose observational content is true. Accordingly you do not need to worry too much about the fact that the coherence measures do not take you to true theories. Instead you can employ the coherence measures to take you to empirically adequate theories. In addition, they favor the theory with more empirical content among two empirically adequate theories after finitely many steps of observation and for every observation thereafter.

However, the achieved results, pleasing as they might be, do not validate the coherentist's twofold claim that (*i*) a theory is a good theory given some observational data *if and only if* that theory coheres with the observational data,[15] and (*ii*) a theory is better than another theory given some observational data *if and only if* the first theory coheres more with the observational data than the second. First, there are other measures, most prominently measures of incremental confirmation, that satisfy the two requirements on adequate measures of

[15] For Olsson's coherence measure this is obvious since Olsson (2002) does not even define the qualitative notion of coherence.

confirmation (Huber, 2005; 2008). Accordingly, it does not follow that we are epistemically obliged to employ the coherence measures for the purpose of theory assessment, since there are some confirmation measures that satisfy both conditions too. Confirmation measures that satisfy both conditions are, for example, Carnap's distance measure d and the Joyce-Christensen measure S (Christensen 1999, Joyce 1999), as has been demonstrated by Huber (2005, 2008).[16] More specifically, one can show that all confirmation measures that assign a maximum degree of confirmation if the observational data implies the theory cannot distinguish between informative and uninformative true theories. Thus according to the presented approach they are unacceptable as measures for the purpose of theory assessment, since they do not even satisfy the minimal conditions put forward in section 2.2. This includes the log-likelihood confirmation measures of Fitelson (2001), Good (1960), and Kemeny-Oppenheim (1952), the normalized distance measures by Crupi et al. (2007), and the normalized log-ratio measure defended by Shogenji (2012) and Atkinson (2012).[17] However, as already said some confirmation measures satisfy the minimal conditions put forward in section. 2.2. Thus, at most, the achieved results show that we are epistemically permitted to use the coherence measures to assess theories since they lead us to believe or accept informative true theories. Alongside these considerations, the latter property of the considered coherence measures constitutes a means-ends justification for the former epistemic permission.

Second, the two requirements put forward in section 2.2 are only minimal conditions on adequate theory assessment functions. We have not specified sufficient conditions for being an adequate measure of confirmation. Hence, we cannot infer that these measures are adequate measures of theory assessment simply from the fact that the coherence measures satisfy the two minimal requirements. This circumstance is aggravated by the fact that at a given point in time the three measures of coherence can give different verdicts with respect to a pair of theories T_1 and T_2. In many cases the available observational data at some point in time might be more coherent with T_1 than with T_2 according to one measure, but the other way around according to another measure.

[16] Not all confirmation measures satisfy this minimal condition though. Huber (2005, 2008) argues that for example the log-likelihood confirmation measure championed by Fitelson (2001), Good (1960), and Kemeny and Oppenheim (1952) is not an adequate measures for the purpose of theory assessment. The reason is that the "log-likelihood ratio measure l neither distinguishes between informative and uninformative true nor between informative and uninformative false theories" (Huber 2005: 1158). This renders it even more interesting that Fitelson's coherence measure distinguishes between informative and uninformative truth. After all, Fitelson's coherence measure, when it is applied to pairwise coherence of theory and observational data, is the average of the two incremental confirmations (in the sense of Kemeny-Oppenheim 1952) between the theory and observational data.

[17] Interestingly, however, the normalized log-ratio measure suggested by Shogenji (2012) and discussed in detail by Atkinson (2012) is motivated by similar considerations as the theory of theory assessment by Huber (2005, 2008). However, a detailed comparison of both approaches has yet to be undertaken.

To sum up, the coherence measures satisfy the two most important criteria for adequate theory assessment functions. First, if we compare two theories, one of which is true and the other false, then the true theory coheres more with the observational data than the false theory (after finitely many steps of observation in a sequence of separating observational statements and for every observation thereafter). This shows that the coherence measures are truth conducive. This is an important and novel result for Bayesian coherentism, especially in the light of the previous impossibility results by Bovens and Hartmann (2003) and Olsson (2005). Second, if we compare two theories, both of which are true but one of which is logically stronger, then the logically stronger theory—i.e., the more informative theory—coheres more with the observational data (after finitely many steps of observation in a sequence of separating observational statements and for every observation thereafter). However, these strong results do not yet show that we are epistemically obliged to use coherence measures to assess theories, yet alone which coherence measure we have to use. There are other measures that satisfy the same conditions and that also lead us to informative true theories. Future investigations into the differences and similarities between the coherentist approach and the confirmation theorist approach will have to determine which approach is more suitable for the purpose of theory assessment.

A Proof of Theorems

Proof of Theorem 3.3: $\forall p\ \forall \epsilon > 0\ \exists \delta_\epsilon > 0 : [p(T_2 \wedge E) > 0\ \&\ 0 < p(E) < 1\ \&\ i_p(T_1, E) \geq i_p(T_2, E) + \epsilon\ \&\ p(T_1|E) \geq p(T_2|E) - \delta_\epsilon] \Rightarrow \mathcal{C}_{O,p}(T_1, E) > \mathcal{C}_{O,p}(T_2, E)$

Proof: Let p be a probability function and let $p(T_2 \wedge E) > 0$ and $0 < p(E) < 1$. Let ϵ be arbitrary with $\epsilon > 0$.

The proof is straightforward if δ_ϵ is chosen arbitrary with $0 < \delta_\epsilon \times p(E) < \frac{p(T_2 \wedge E)}{1-p(\neg T_2 \wedge \neg E)} \times [\epsilon \times p(\neg E)]$. We know there is such an δ_ϵ, because $p(T_2 \wedge E) > 0$, $0 < p(E) < 1$ and $\epsilon > 0$.

We have to prove that: $i_p(T_1, E) \geq i_p(T_2, E) + \epsilon \Rightarrow [p(T_1|E) \geq p(T_2|E) - \delta_\epsilon \Rightarrow \mathcal{C}_{O,p}(T_1, E) > \mathcal{C}_{O,p}(T_2, E)]$

2.) Suppose $i_p(T_1, E) \geq i_p(T_2, E) + \epsilon$, this means $p(\neg T_1 \wedge \neg E) \geq p(\neg T_2 \wedge \neg E) + \epsilon \times p(\neg E)$ which implies $1 - p(\neg T_1 \wedge \neg E) \leq 1 - [p(\neg T_2 \wedge \neg E) + \epsilon \times p(\neg E)]$

3.) Suppose $p(T_1|E) \geq p(T_2|E) - \delta_\epsilon$, i.e. $p(T_1 \wedge E) \geq p(T_2 \wedge E) - \delta_\epsilon \times p(E)$

4.) $\frac{\delta_\epsilon \times p(E)}{p(T_2 \wedge E)} < \frac{\epsilon \times p(\neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 1.)

5.) $[1 - \frac{\delta_\epsilon \times p(E)}{p(T_2 \wedge E)}] > 1 - \frac{\epsilon \times p(\neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 4.)

6.) $\frac{p(T_2 \wedge E)}{p(T_2 \wedge E)} - \frac{\delta_\epsilon \times p(E)}{p(T_2 \wedge E)} > 1 - \frac{\epsilon \times p(\neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 5.)

7.) $\frac{p(T_2 \wedge E) - \delta_\epsilon \times p(E)}{p(T_2 \wedge E)} > 1 - \frac{\epsilon \times p(\neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 6.)

8.) $\frac{p(T_1 \wedge E)}{p(T_2 \wedge E)} > 1 - \frac{\epsilon \times p(\neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 3. and 7.)

9.) $\frac{p(T_1 \wedge E)}{p(T_2 \wedge E)} > \frac{1-p(\neg T_2 \wedge \neg E)}{1-p(\neg T_2 \wedge \neg E)} - \frac{\epsilon \times p(\neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 8.)

10.) $\frac{p(T_1 \wedge E)}{p(T_2 \wedge E)} > \frac{[1-p(\neg T_2 \wedge \neg E)] - \epsilon \times p(\neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 9.)

11.) $\frac{p(T_1 \wedge E)}{p(T_2 \wedge E)} > \frac{1-[p(\neg T_2 \wedge \neg E) + \epsilon \times p(\neg E)]}{1-p(\neg T_2 \wedge \neg E)}$ (from 10.)

12.) $\frac{p(T_1 \wedge E)}{p(T_2 \wedge E)} > \frac{1-p(\neg T_1 \wedge \neg E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 2. and 11.)

13.) $\frac{p(T_1 \wedge E)}{1-p(\neg T_1 \wedge \neg E)} > \frac{p(T_2 \wedge E)}{1-p(\neg T_2 \wedge \neg E)}$ (from 12.)

14.) $\frac{p(T_1 \wedge E)}{p(T_1 \vee E)} > \frac{p(T_2 \wedge E)}{p(T_2 \vee E)}$ (from 13.)

15.) $\mathcal{C}_{O,p}(T_1, E) > \mathcal{C}_{O,p}(T_2, E)$ (from 14. and the definition of $\mathcal{C}_{O,p}$)

Proof of Theorem 3.6: $\forall p\ \forall \epsilon > 0\ \exists \delta_\epsilon > 0 : [cont_p(T_1) \geq cont_p(T_2) + \epsilon\ \&\ 0 \neq p(T_1|E) \geq p(T_2|E) - \delta_\epsilon] \Rightarrow \mathcal{C}_{S,p}(T_1, E) > \mathcal{C}_{S,p}(T_2, E)$

Proof: Let p be a probability function with $p(E \wedge T_1) > 0$ and let $\epsilon > 0$.

The proof is straightforward if δ_ϵ is chosen arbitrary satisfying $0 < \delta_\epsilon < \frac{p(T_1 \wedge E)}{p(E) \times p(T_1)} \times \epsilon$. We know there is such an $\delta_\epsilon \in \mathbb{R}$, because $p(T_1 \wedge E) > 0$.

Proof of Theorem 3.10: $\forall p\ \forall \epsilon > 0\ \exists \delta_\epsilon > 0 : [p(T_1 \wedge E) > 0\ \&\ cont_p(T_1) \geq cont_p(T_2) + \epsilon\ \&\ p(T_1|E) \geq p(T_2|E) - \delta_\epsilon] \Rightarrow \mathcal{C}_{F,p}(T_1, E) > \mathcal{C}_{F,p}(T_2, E)$

Proof: Let p be an arbitrary probability function with $p(T_1 \wedge E) > 0$ and let $\epsilon > 0$.

The proof is straightforward if δ_ϵ chosen arbitrary satifying $0 < \delta_\epsilon \leq \frac{p(T_1 \wedge E)}{p(E) \times p(T_1)} \times \epsilon$. We know that there is such an δ_ϵ because $p(T_1 \wedge E) > 0$ and $\epsilon > 0$.

Proof of Theorem 4.1 for Olsson's coherence measure: Let $e_1, \ldots, e_n, \ldots$ be a sequence of sentences of $\mathcal{L}$ which separates $Mod_{\mathcal{L}}$, and let $e_i^w = e_i$, if $w \vDash e_i$ and $\neg e_i$ otherwise. Let p be a strict (or regular) probability function on $\mathcal{L}$. Let p^* be the unique determined probability function on the smallest σ-field $\mathcal{A}$ containing the field $\{Mod(A) : A \in \mathcal{L}\}$ satisfying $p^*(Mod(A)) = p(A)$ for all $A \in \mathcal{L}$, where $Mod(A) = \{w \in Mod_{\mathcal{L}} : w \vDash A\}$ and $Mod_{\mathcal{L}}$ is the set of all maximal-consistent sets of sentences of $\mathcal{L}$ including instances.
Then according to the Gaifman-Snir Theorem (Gaifman and Snir 1982) there is a $X \subseteq Mod_{\mathcal{L}}$ with $p^*(X) = 1$, such that the following holds for every $w \in X$ and all theories T_1 and T_2 of $\mathcal{L}$.

$$lim_{n\to\infty} p(T_1|E_n^w) = \mathcal{I}(T_1, w)$$

where $\mathcal{I}(T_1, w) = 1$, if $w \vDash T_1$ and 0 otherwise.

1.) Suppose additionally that $w' \in X$ and $w' \vDash T_1$ and $w' \vDash \neg T_2$.
By the Gaifman-Snir (1982) Theorem and the assumptions the following holds:

$lim_{n\to\infty}[p(T_1|E_n^{w'})] = 1$ and $lim_{n\to\infty}[p(T_2|E_n^{w'})] = 0$.

And since the following holds true:

$lim_{n\to\infty}[p(T_1 \vee E_n^{w'})] \in [p(T_1), 1]$ and $lim_{n\to\infty}[p(T_2 \vee E_n^{w'})] \in [p(T_2), 1]$

we can conclude that

$$lim_{n\to\infty}[p(T_1|E_n^{w'})] > lim_{n\to\infty}[p(T_2|E_n^{w'}) \times \frac{p(T_1 \vee E_n^{w'})}{p(T_2 \vee E_n^{w'})}].$$

This implies $\exists n \forall m \geq n : p(T_1|E_m^{w'}) > p(T_2|E_m^{w'}) \times \frac{p(T_1 \vee E_m^{w'})}{p(T_2 \vee E_m^{w'})}$.

From which we can infer that: $\exists n \forall m \geq n : \frac{p(T_1 \wedge E_m^{w'})}{p(T_1 \vee E_m^{w'})} > \frac{p(T_2 \wedge E_m^{w'})}{p(T_2 \vee E_m^{w'})}$.

which implies that: $\exists n \forall m \geq n :$

$$\mathcal{C}_{O,p}(T_1, E_m^{w'}) > \mathcal{C}_{O,p}(T_2, E_m^{w'})$$

2.) Additionally assume that $w' \in X$ and $w' \vDash T_1 \wedge T_2$ and $T_1 \vdash T_2$ but $T_2 \nvdash T_1$. Since p is by assumption a strict probability function is must hold that $cont_p(T_1) > cont_p(T_2)$ and hence that $p(T_2) > p(T_1)$.

By the Gaifman-Snir Theorem we can infer that:

$$lim_{m\to\infty}p(T_1|E_m^{w'}) = lim_{m\to\infty}p(T_2|E_m^{w'}) = 1$$

with this we get that for all ϵ there exists a n_ϵ such that for all $m \geq n_\epsilon$ holds that:

$$|p(T_2|E_m^{w'}) - p(T_1|E_m^{w'})| < \epsilon$$

Then define $\epsilon = \frac{[p(T_2)-p(T_1)]}{2}$

From this definition we get that for all $m \geq n_\epsilon$:

$p(\neg T_1|\neg E_m^{w'}) - p(\neg T_2|\neg E_m^{w'}) > \epsilon$ since:

1. $p(T_2|E_m^{w'}) - p(T_1|E_m^{w'}) < \epsilon$ for all $m \geq n_\epsilon$ as assumed above.

2. $[p(T_2) - p(T_1)] - [p(T_2|E_m^{w'}) - p(T_1|E_m^{w'})] > [p(T_2) - p(T_1)] - \epsilon$

3.) $[p(T_2) - p(T_1)] - [p(T_2|E_m^{w'}) - p(T_1|E_m^{w'})] > \epsilon$

4.) $[p(T_2) - p(T_1)] - [p(T_2|E_m^{w'}) - p(T_1|E_m^{w'})] \times p(E_m^{w'}) > \epsilon$

(from 3.)

5.) $[p(T_2) - p(T_1)] - [p(T_2 \wedge E_m^{w'}) - p(T_1 \wedge E_m^{w'})] > \epsilon$

6.) $[p(T_2) - p(T_2 \wedge E_m^{w'})] - [p(T_1) - p(T_1 \wedge E_m^{w'})] > \epsilon$

7.) $[p(T_2) + p(E_m^{w'}) - p(T_2 \wedge E_m^{w'})] - [p(T_1) + p(E_m^{w'}) - p(T_1 \wedge E_m^{w'})] > \epsilon$

8.) $[p(T_2 \vee E_m^{w'})] - [p(T_1 \vee E_m^{w'})] > \epsilon$

9.) $[1 - p(T_1 \vee E_m^{w'})] - [1 - p(T_2 \vee E_m^{w'})] > \epsilon$

10.) $p(\neg T_1 \wedge \neg E_m^{w'}) - p(\neg T_2 \wedge \neg E_m^{w'}) > \epsilon$

11.) $p(\neg T_1 \wedge \neg E_m^{w'}) - p(\neg T_2 \wedge \neg E_m^{w'}) > \epsilon$

12.) $p(\neg T_1|\neg E_m^{w'}) - p(\neg T_2|\neg E_m^{w'}) > \frac{\epsilon}{p(\neg E_m^{w'})} > \epsilon$

From which we can infer that:

(I) $\forall m \geq n_\epsilon p(\neg T_1|\neg E_m^{w'}) > p(\neg T_2|\neg E_m^{w'}) + \epsilon$

By Theorem 3.3 we can infer that there is a δ_ϵ such that:

$$p(T_1|E_m^{w'}) \geq p(T_2|E_m^{w'}) - \delta_\epsilon \Rightarrow \mathcal{C}_{O,p}(T_1, E_m^{w'}) > \mathcal{C}_{O,p}(T_2, E_m^{w'})$$

Again by the Gaifman-Snir Theorem we get that there is a n_{δ_ϵ} such that for all $m \geq n_{\delta_\epsilon}$

$$p(T_2|E_m^{w'}) - p(T_1|E_m^{w'}) < \delta_\epsilon$$

or equivalently: (II) $p(T_1|E_m^{w'}) > p(T_2|E_m^{w'}) - \delta_{n_1}$

Again by Theorem 3.3 and (I) and (II) we can conclude: For all $m \geq n^*$, where n^* =max.$\{n_{\delta_\epsilon}, n_\epsilon\}$ it holds that:

$$\mathcal{C}_{O,p}(T_1, E_m^{w'}) > \mathcal{C}_{O,p}(T_2, E_m^{w'})$$

Which implies that there a n such that for all For all $m \geq n$:

$$\mathcal{C}_{O,p}(T_1, E_m^{w'}) > \mathcal{C}_{O,p}(T_2, E_m^{w'})$$

Proof of Theorem 4.1 for Shogenji's coherence measure: Let $e_1, \ldots, e_n, \ldots$ be a sequence of sentences of $\mathcal{L}$ which separates $Mod_\mathcal{L}$, and let $e_i^w = e_i$, if $w \vDash e_i$ and $\neg e_i$ otherwise. Let p be a strict (or regular) probability function on $\mathcal{L}$. Let p^* be the unique determined probability function on the smallest σ-field $\mathcal{A}$ containing the field $\{Mod(A) : A \in \mathcal{L}\}$ satisfying $p^*(Mod(A)) = p(A)$ for all $A \in \mathcal{L}$, where $Mod(A) = \{w \in Mod_\mathcal{L} : w \vDash A\}$ and $Mod_\mathcal{L}$ is the set of all maximal-consistent sets of sentences of $\mathcal{L}$ including instances.
Then according to the Gaifman-Snir Theorm (Gaifman and Snir 1982) there is a $X \subseteq Mod_\mathcal{L}$ with $p^*(X) = 1$, such that the following holds for every $w \in X$ and all theories T_1 and T_2 of $\mathcal{L}$:

$$lim_{n\to\infty} p(T_1|E_n^w) = \mathcal{I}(T_1, w)$$

where $\mathcal{I}(T_1, w) = 1$, if $w \vDash T_1$ and 0 otherwise.

1.) Suppose additionally that $w' \in X$, $w' \vDash T_1$ and $w' \vDash \neg T_2$.

We know that $lim_{n\to\infty}[p(T_1|E_n^{w'})] = 1$ and $lim_{n\to\infty}[p(T_2|E_n^{w'})] = 0$ by the Gaifman-Snir Theorem.

So we can infer that:
$lim_{n\to\infty}[p(T_1|E_n^{w'}) \times \frac{1}{p(T_1)}] = \frac{1}{p(T_1)} > 1 > lim_{n\to\infty}[p(T_2|E_n^{w'}) \times \frac{1}{p(T_2)}] = 0$

Let $\epsilon = \frac{\frac{1}{p(T_1)}}{2}$. Then it holds that: $\exists n \forall m \geq n : |\frac{1}{p(T_1)} - \mathcal{C}_{S,p}(T_1, E_m^{w'})| < \epsilon$ and $\mathcal{C}_{S,p}(T_1, E_m^{w'}) > 1$ and $\exists n' \forall m \geq n' : |0 - \mathcal{C}_{S,p}(T_2, E_m^{w'})| < \epsilon$ and $\mathcal{C}_{S,p}(T_2, E_m^{w'}) < 1$.

Now let n_1 =max.$\{n, n'\}$. Then it holds for all $m \geq n_1$:

$$\mathcal{C}_{S,p}(T_1, E_m^{w'}) > 1 > \mathcal{C}_{S,p}(T_2, E_m^{w'})$$

2.) Additionally assume that $w' \in X$ and $w' \vDash T_1 \wedge T_2$ and $T_1 \vdash T_2$ but $T_2 \nvdash T_1$. Since p is by assumption a strict probability function is must hold that $cont_p(T_1) > cont_p(T_2)$ and hence that $p(T_2) > p(T_1)$.

Thereby we know that: $lim_{n\to\infty}[p(T_1|E_n^{w'})] = 1$ and $lim_{n\to\infty}[p(T_2|E_n^{w'})] = 1$ by the Gaifman-Snir Theorem and the assumption that $w' \vDash T_1 \wedge T_2$.

This implies that $lim_{n\to\infty}[p(T_1|E_n^{w'}) \times \frac{1}{p(T_1)}] = \frac{1}{p(T_1)} > lim_{n\to\infty}[p(T_2|E_n^{w}) \times \frac{1}{p(T_2)}] = \frac{1}{p(T_2)}$

Since $cont_p(T_1) > cont_p(T_2)$ we can infer that $p(T_1) < p(T_2)$ and $\frac{1}{p(T_1)} > \frac{1}{p(T_2)}$
Now let $\epsilon = \frac{\frac{1}{p(T_1)} - \frac{1}{p(T_2)}}{2}$. Then it holds that: $\exists n \forall m \geq n : |\frac{1}{p(T_1)} - \mathcal{C}_{S,p}(T_1, E_m^{w})| < \epsilon$ and $\exists n' \forall m \geq n' |\frac{1}{p(T_2)} - \mathcal{C}_{S,p}(T_2, E_m^{w})| < \epsilon$.

Now let n_1 =max.$\{n, n'\}$. Then it holds for all $m \geq n_1$:

$$\mathcal{C}_{S,p}(T_1, E_m^{w}) > \mathcal{C}_{S,p}(T_2, E_m^{w})$$

Proof of Theorem 4.1 for Fitelson's coherence measure: Let $e_1, \ldots, e_n, \ldots$ be a sequence of sentences of $\mathcal{L}$ which separates $Mod_{\mathcal{L}}$, and let $e_i^w = e_i$, if $w \vDash e_i$ and $\neg e_i$ otherwise. Let p be a strict (or regular) probability function on $\mathcal{L}$. Let p^* be the unique determined probability function on the smallest σ-field $\mathcal{A}$ containing the field $\{Mod(A) : A \in \mathcal{L}\}$ satisfying

$p^*(Mod(A)) = p(A)$ for all $A \in \mathcal{L}$, where $Mod(A) = \{w \in Mod_{\mathcal{L}} : w \vDash A\}$ and $Mod_{\mathcal{L}}$ is the set of all maximal-consistent sets of sentences of $\mathcal{L}$ including instances.
Then according to the Gaifman-Snir Theorm (Gaifman and Snir 1982) there is a $X \subseteq Mod_{\mathcal{L}}$ with $p^*(X) = 1$, such that the following holds for every $w \in X$ and all theories T_1 and T_2 of $\mathcal{L}$.

$$lim_{n\to\infty} p(T_1|E_n^w) = \mathcal{I}(T_1, w)$$

where $\mathcal{I}(T_1, w) = 1$, if $w \vDash T_1$ and 0 otherwise.

1.) Assume additionally that $w' \in X$ and and $w' \vDash T_1$ and $w' \vDash \neg T_2$.
First note that:

$$lim_{n\to\infty}[\mathcal{C}_{F,p}(T_i, E_n^{w'})] =$$

$$\frac{1}{2}[lim_{n\to\infty}[\frac{p(T_i|E_n^{w'}) - p(T_i|\neg E_n^{w'})}{p(T_i|E_n^{w'}) + p(T_i|\neg E_n^{w'})} \times \frac{\frac{1}{p(T_i)}}{\frac{1}{p(T_i)}}]+$$

$$lim_{n\to\infty}[\frac{p(E_n^{w'}|T_i) - p(E_n^{w'}|\neg T_i)}{p(E_n^{w'}|T_i) + p(E_n^{w'}|\neg T_i)} \times \frac{\frac{1}{p(E_n^{w'})}}{\frac{1}{p(E_n^{w'})}}]] =$$

$$\frac{1}{2}[lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_i, E_n^{w'}) - \mathcal{C}_{S,p}(T_i, \neg E_n^{w'})}{\mathcal{C}_{S,p}(T_i, E_n^{w'}) + \mathcal{C}_{S,p}(T_i, \neg E_n^{w'})}]+$$

$$lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_i, E_n^{w'}) - \mathcal{C}_{S,p}(\neg T_i, E_n^{w'})}{\mathcal{C}_{S,p}(T_i, E_n^{w'}) + \mathcal{C}_{S,p}(\neg T_i, E_n^{w'})}]]$$

and since we know that under the above assumptions by Theorem 4.1 the following holds:

1. $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_1, E_n^{w'})] = \frac{1}{p(T_1)}$
2. $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_1, \neg E_n^{w'})] \leq \frac{1}{p(T_1)}$
3. $lim_{n\to\infty}[\mathcal{C}_{S,p}(\neg T_1, E_n^{w'})] = 0$
4. $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_2, E_n^{w'})] = 0$
5. $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_2, \neg E_n^{w'})] \geq 0$
6. $lim_{n\to\infty}[\mathcal{C}_{S,p}(\neg T_2, E_n^{w'})] = \frac{1}{p(\neg T_2)}$

we can conclude that:

$$lim_{n\to\infty}[\mathcal{C}_{F,p}(T_1, E_n^{w'})] > 0 > lim_{n\to\infty}[\mathcal{C}_{F,p}(T_2, E_n^{w'})$$

Which implies that:

$$\exists n \forall m \geq n : \mathcal{C}_{F,p}(T_1, E_m^{w'}) > 0 > \mathcal{C}_{F,p}(T_2, E_m^{w'})$$

2.) Additionally assume that $w' \in X$ and $w' \vDash T_1 \wedge T_2$ and $T_1 \vdash T_2$ but $T_2 \nvdash T_1$. Since p is by assumption a strict probability function is must hold that $cont_p(T_1) > cont_p(T_2)$ and hence that $p(T_2) > p(T_1)$.

We already know that:

$$lim_{n\to\infty}[\mathcal{C}_{F,p}(T_i, E_n^{w'})] =$$
$$\frac{1}{2}[lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_i, E_n^{w'}) - \mathcal{C}_{S,p}(T_i, \neg E_n^{w'})}{\mathcal{C}_{S,p}(T_i, E_n^{w'}) + \mathcal{C}_{S,p}(T_i, \neg E_n^{w'})}]+$$
$$lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_i, E_n^{w'}) - \mathcal{C}_{S,p}(\neg T_i, E_n^{w'})}{\mathcal{C}_{S,p}(T_i, E_n^{w'}) + \mathcal{C}_{S,p}(\neg T_i, E_n^{w'})}]]$$

and since we know that under the above assumptions by Theorem 4.1 the following holds:

1. $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_1, E_n^{w'})] = \frac{1}{p(T_1)}$
2. $lim_{n\to\infty}[\mathcal{C}_{S,p}(\neg T_1, E_n^{w'})] = 0$
3. $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_2, E_n^{w'})] = \frac{1}{p(T_2)}$
4. $lim_{n\to\infty}[\mathcal{C}_{S,p}(\neg T_2, E_n^{w'})] = 0$

Accordingly, in order to prove Theorem 4.3 it is sufficient to prove that:

$$lim_{n\to\infty}[\mathcal{C}_{S,p}(T_2, \neg E_n^{w'})] \geq lim_{n\to\infty}[\mathcal{C}_{S,p}(T_1, \neg E_n^{w'})]$$

We already know that:

$$lim_{n\to\infty}[\mathcal{C}_{S,p}(T_1, E_n^{w'})] = \frac{1}{p(T_1)} > lim_{n\to\infty}[\mathcal{C}_{S,p}(T_2, E_n^{w'})] = \frac{1}{p(T_2)} \text{ (since } p(T_1) < p(T_2))$$

This implies

$$lim_{n\to\infty}[p(E_n^{w'}|T_1)] > lim_{n\to\infty}[p(E_n^{w'}|T_2)] \Leftrightarrow$$
$$lim_{n\to\infty}[1 - p(E_n^{w'}|T_2)] > lim_{n\to\infty}[1 - p(E_n^{w'}|T_1)] \Leftrightarrow$$
$$lim_{n\to\infty}[p(\neg E_n^{w'}|T_2)] > lim_{n\to\infty}[p(\neg E_n^{w'}|T_1)] \Leftrightarrow$$
$$lim_{n\to\infty}[p(\neg E_n^{w'}|T_2) \times \frac{1}{p(\neg E_n^{w'})}] > lim_{n\to\infty}[p(\neg E_n^{w'}|T_1) \times \frac{1}{p(\neg E_n^{w'})}] \Leftrightarrow$$
$$lim_{n\to\infty}[\mathcal{C}_{S,p}(T_2, \neg E_n^{w'})] \geq lim_{n\to\infty}[\mathcal{C}_{S,p}(T_1, \neg E_n^{w'})]$$

We can conclude

$$lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_1, E_n^{w'}) - \mathcal{C}_{S,p}(T_1, \neg E_n^{w'})}{\mathcal{C}_{S,p}(T_1, E_n^{w'}) + \mathcal{C}_{S,p}(T_1, \neg E_n^{w'})}] > lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_2, E_n^{w'}) - \mathcal{C}_{S,p}(T_2, \neg E_n^{w'})}{\mathcal{C}_{S,p}(T_2, E_n^{w'}) + \mathcal{C}_{S,p}(T_2, \neg E_n^{w'})}]$$

and

$$lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_2, E_n^{w'}) - \mathcal{C}_{S,p}(\neg T_2, E_n^{w'})}{\mathcal{C}_{S,p}(T_2, E_n^{w'}) + \mathcal{C}_{S,p}(\neg T_2, E_n^{w'})}] = lim_{n\to\infty}[\frac{\mathcal{C}_{S,p}(T_2, E_n^{w'}) - \mathcal{C}_{S,p}(\neg T_2, E_n^{w'})}{\mathcal{C}_{S,p}(T_2, E_n^{w'}) + \mathcal{C}_{S,p}(\neg T_2, E_n^{w'})}]$$

Which proves that:

$$\exists n \forall m \geq n : \mathcal{C}_{F,p}(T_1, E_m^{w'}) > \mathcal{C}_{F,p}(T_2, E_m^{w'})$$

Proof of Theorem 4.2: Let $e_1,\ldots,e_n,\ldots$ be a sequence of statements of $\mathcal{L}$ which separates $Mod_{\mathcal{L}}$, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\neg e_i$ otherwise. Let p be a strict (or regular) probability function on $\mathcal{L}$. Let p^* be the unique probability function on the smallest σ-field $\mathcal{A}$ containing the field $\{Mod(A) : A \in \mathcal{L}\}$ satisfying $p^*(Mod(A)) = p(A)$ for all $A \in \mathcal{L}$, where $Mod(A) = \{w \in Mod_{\mathcal{L}} : w \vDash A\}$ and $Mod_{\mathcal{L}}$ is the set of all maximal-consistent sets of statements of $\mathcal{L}$ including instances.
Then there is an $X \subseteq Mod_{\mathcal{L}}$ with $p^*(X) = 1$ such that the following holds for every $w \in X$ and all theories T_1 and T_2 of $\mathcal{L}$.

$$lim_{n\to\infty} p(T_1|E_n^w) = \mathcal{I}(T_1, w)$$

where $\mathcal{I}(T_1, w) = 1$, if $w \vDash T_1$ and 0 otherwise.

Hence, if $w \vDash \neg T_i$, then:

$$lim_{n\to\infty} p(T_i|E_n^w) = 0$$

This implies that
$lim_{n\to\infty}[\mathcal{C}_{O,p}(T_i, E_m^w)] = 0$ since $0 \leq \mathcal{C}_{O,p}(T_i, E_n^w) \leq p(T_i|E_n^w)$, for all n.

Furthermore, $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_i, E_m^w)] = 0$, since $\mathcal{C}_{S,p}(T_i, E_n^w) = \dfrac{p(T_i|E_n^w)}{p(T_1)}$, for all n.

In addition, we know from proof of theorem 4.1 for Fitelson's measure $\mathcal{C}_{F,p}$ that

$$lim_{n\to\infty}[\mathcal{C}_{F,p}(T_i, E_n^{w'})] =$$

$$\frac{1}{2} lim_{n\to\infty} \left[\frac{\mathcal{C}_{S,p}(T_i, E_n^{w'}) - \mathcal{C}_{S,p}(T_i, \neg E_n^{w'})}{\mathcal{C}_{S,p}(T_i, E_n^{w'}) + \mathcal{C}_{S,p}(T_i, \neg E_n^{w'})} \right] +$$

$$\frac{1}{2} lim_{n\to\infty} \left[\frac{\mathcal{C}_{S,p}(T_i, E_n^{w'}) - \mathcal{C}_{S,p}(\neg T_i, E_n^{w'})}{\mathcal{C}_{S,p}(T_i, E_n^{w'}) + \mathcal{C}_{S,p}(\neg T_i, E_n^{w'})} \right]$$

We already said that $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_i, E_m^w)] = 0$ and $lim_{n\to\infty}[\mathcal{C}_{S,p}(\neg T_i, E_m^w)] = 1$.

Furthermore, $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_i, E_m^w)] = 0$ iff $lim_{n\to\infty}[p(T_i \wedge E_m^w)] = 0$ and hence, $lim_{n\to\infty}[p(T_i \wedge \neg E_m^w)] = p(T_i)$. Thus $lim_{n\to\infty}[\mathcal{C}_{S,p}(T_i, \neg E_m^w)] = \dfrac{1}{lim_{n\to\infty}[p(\neg E_m^w)]}$, where $lim_{n\to\infty}[p(\neg E_m^w)] > 0$. Hence,

$$\frac{1}{2} \left[\frac{0 - lim_{n\to\infty}[p(\neg E_n^{w'})]}{0 + lim_{n\to\infty}[p(\neg E_n^{w'})]} \right] + \left[\frac{0-1}{0+1} \right] = -1$$

References

[1] Atkinson, D. (2012). Confirmation and Justification. A Commentary on Shogenji's Measure. *Synthese* 184, 49–61.

[2] Bovens, L., Hartmann, S. (2003). *Bayesian Epistemology.* Oxford: Oxford University Press.

[3] Brössel, P. (forthcoming). Correlation and Truth. in V. Karakostas, D. Dieks (eds.). *Recent Progress in Philosophy of Science: Perspectives and Foundational Problems.* Dordrecht: Springer.

[4] Christensen, D. (1999). Measuring Confirmation. *Journal of Philosophy* 96, 437–461.

[5] Carnap, R., Bar-Hillel, Y. (1952). *An outline of a theory of semantic information.* Tech. Report no. 247 Research Laboratory of Electronics, Massachusetts Institute of Technology.

[6] Douven, I., Meijs, W. (2007). Measuring Coherence. *Synthese* 156, 405–425.

[7] Earman, J. (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory.* Cambridge, Massachusetts: MIT Press.

[8] Fitelson, B. (2001). *Studies in Bayesian Confirmation Theory.* PhD. Dissertation, University of Wisconsin-Madison (Philosophy).

[9] Fitelson, B. (2003). A Probabilistic Theory of Coherence. *Analysis* 63, 194–199.

[10] Glass, D. (2007). Coherence measures and inference to the best explanation, *Synthese* 157, 275–296 .

[11] Gaifman, H., Snir, M. (1982). Probabilities over rich languages, testing, and randomness. *Journal of Symbolic Logic* 47, 495–548.

[12] Hawthorne, J. (2011), Confirmation Theory. in P. S. Bandyopadhyay, M. R. Forster (eds.), *Handbook of Philosophy of Science, Volume 7: Philosophy of Statistics*. Amsterdam: Elsevier.

[13] Good, I. J. (1960). Weight of Evidence, Corroboration, Explanatory Power, Information and the Utility of Experiments. *Journal of the Royal Statistical Society.* Series B (Methodological) 22, 319–331.

[14] Hempel, C. G. (1960). Inductive Inconsistencies. *Synthese* 12, 439–469.

[15] Hempel, C. G., Oppenheim, P. (1948). Studies in the Logic of Explanation. in *Aspects of Scientific Explanation and other Essays in the Philosophie of Science.* (1965), New York: Free Press, 245–291.

[16] Hilpinen; R. (1970). On the Information Provided by Observations. in J. Hintikka, P. Suppes (eds.), *Information and Inference*. Dordrecht: D. Reidel, 97–123.

[17] Hintikka, J. (1999). The Varieties of Information and Scientific Explanation. in *Jaakko Hintikka Selected Papers, Volume 5: Inquiry as Inquiry: A Logic of Scientific Discovery*, Dordrecht: Kluwer Academic Publishers.

[18] Hintikka, J., Pietarinen, J. (1966). Semantic Information and Inductive Logic. in J. Hintikka, P. Suppes (eds.), *Aspects of Inductive Logic*, Amsterdam: North-Holland Publ. Comp, 96–112.

[19] Huber, F. (2008). Assessing Theories, Bayes Style. *Synthese* 161, 89–118.

[20] Huber, F. (2005). What Is the Point of Confirmation? *Philosophy of Science* 72, 1146–1159.

[21] Joyce, J. (1999). *Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

[22] Kemeny, J., Oppenheim, P. (1952). Degrees of Factual Support. *Philosophy of Science* 19, 307–324.

[23] Levi, I. (1961). Decision Theory and Confirmation. *The Journal of Philosophy* 58, 614–625.

[24] Levi, I. (1967). *Gambling with Truth*. New York: A. A. Knopf.

[25] Miller, D. (1974). Popper's Qualitative Theory of Verisimilitude. *British Journal for the Philosophy of Science* 25, 166–177.

[26] Olsson, E. (2002). What Is the Problem of Coherence and Truth? *Journal of Philosophy* 94, 246–272.

[27] Olsson, E. (2005). *Against Coherence: Truth, Probability, and Justifiction*. Oxford: Oxford University Press.

[28] Popper, K. (1954). Degree of Confirmation. *The British Journal of the Philosophy of Science* 5, 143–149.

[29] Popper, K. (1968). *The Logic of Scientific Discovery*. London: Hutchinson.

[30] Salmon, W. (2001). Reflections of a Bashful Bayesian: A Reply to Peter Lipton. in G. Hon, S. Rakover (eds.) (2001), *Explanation: Theoretical Approaches and Applications*. Dordrecht: Kluwer, 121–136.

[31] Shogenji, T. (2012). The Degree of Epistemic Justification and the Conjunction Fallacy. *Synthese* 184, 29-.48.

[32] Shogenji, T. (1999). Is Coherence Truth Conducive? *Analysis* 59, 338–345.

[33] Tichý, P. (1974). On Popper's Definitions of Verisimilitude. *British Journal for the Philosophy of Science* 25, 155–160.

[34] Van Fraassen, B. C. (1980). *The Scientific Image*. Oxford: Oxford University Press.