Keynes's Coefficient of Dependence Revisited*

Peter Brössel
Department of Philosophy and
Center for Mind, Brain, and Cognitive Evolution
Ruhr-University Bochum
Universitätsstrasse 150
44780 Bochum
peter.broessel@rub.de

Contents

*The present paper has its origin in a short paper that appeared as "Correlation and Truth" in: V. Karakostas, D. Dieks (eds.). *EPSA11 Recent Progress in Philosophy of Science: Perspectives and Foundational Problems*. Dordrecht: Springer, 41-54. For their helpful remarks and suggestions on previous versions of the paper I am very thankful to Ralf Busse, Franz Huber, Gregory Wheeler, two anonymous referees of this journal, and (especially) Anna-Maria A. Eder and Ben Young.

Abstract

Probabilistic dependence and independence are among the key concepts of Bayesian epistemology. This paper focuses on the study of one specific quantitative notion of probabilistic dependence. More specifically, section 1 introduces Keynes's coefficient of dependence and shows how it is related to pivotal aspects of scientific reasoning such as confirmation, coherence, the explanatory and unificatory power of theories, and the diversity of evidence. The intimate connection between Keynes's coefficient of dependence and scientific reasoning raises the question of how Keynes's coefficient of dependence is related to truth, and how it can be made fruitful for epistemological considerations. This question is answered in section 2 of the paper. Section 3 outlines the consequences the results have for epistemology and the philosophy of science from a Bayesian point of view.

1 Keynes's Coefficient of Dependence and its Relation to Key Concepts in Bayesian Epistemology

1.1 Introduction

Two propositions A and B are probabilistically dependent relative to a probability function Pr if and only if $\Pr(A \cap B) \neq \Pr(A) \times \Pr(B)$; alternatively if and only if $\Pr(B) > 0$ and $\Pr(A|B) \neq \Pr(A)$. Two propositions A and B are probabilistically independent relative to a probability function Pr if and only if $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$.

These qualitative concepts of probabilistic dependence and probabilistic independence are among the key concepts of Bayesian philosophy of science and epistemology. This can be seen most clearly in the following example. One central research question of Bayesian philosophy is: How should we update our degrees of belief conceived of as probabilities? The standard Bayesian rule for updating degree of belief is *Strict Conditionalization*. According to *Strict Conditionalization*, if an agent learns with certainty that B is true (and $\Pr_{old}(B) > 0$), the new degree of belief in A ($\Pr_{new}(A)$) is the old degree of belief in A given B ($\Pr_{old}(A|B)$). Hence, whether $\Pr_{new}(A) \neq \Pr_{old}(A)$ depends on whether $\Pr_{old}(A|B) \neq \Pr_{old}(A)$ and thus on whether A and B are probabilistic dependent or independent with respect to $\Pr_{old}$. If A and B are positively probabilistic dependent (i.e. if $\Pr_{old}(A \cap B) > \Pr_{old}(A) \times \Pr_{old}(B)$), learning B results in an increase in degree of belief in proposition A. If A and B are negatively probabilistic dependent (i.e. if $\Pr_{old}(A \cap B) < \Pr_{old}(A) \times \Pr_{old}(B)$) learning B results in a decrease in degree of belief in proposition A. If A and B are probabilistic independent (i.e. if $\Pr_{old}(A \cap B) = \Pr_{old}(A) \times \Pr_{old}(B)$), learning B does not change one's degree of belief in proposition A.

This paper focuses on the study of probabilistic dependence and independence. More specifically, it focuses on one particularly simple measure of probabilistic dependence, i.e. a quantitative concept of deviation from independence, which was introduced by John Maynard Keynes (1921) and is highly significant for the philosophy of science and epistemology. It is highly significant because one can show that Keynes's *coefficient of dependence is a key element for furthering our understanding of almost any Bayesian attempt to capture or explicate any interesting aspect of scientific reasoning in terms of probabilities. Subsections 1.2–1.5* show how Keynes's coefficient of dependence is related to many aspects of scientific reasoning: confirmation, coherence, explanatory and unificatory power, and diversity of evidence. The intimate connection between Keynes's coefficient of dependence and scientific reasoning raises the question of how this coefficient of dependence is related to truth, and how it can be made fruitful for epistemology. This question is answered in section 2. Section 3 outlines the consequences the results have for the philosophy of science and epistemology from a Bayesian point of view.

1.2 Measuring Probabilistic Dependence: Keynes's Coefficient of Dependence

In contrast to concepts such as confirmation, coherence, explanatory and unificatory power, the concept of probabilistic dependence, i.e. deviation from independence, is not a key concept within traditional approaches to scientific and/or rational reasoning in philosophy of science and epistemology. However, the aim of Bayesian philosophy of science and epistemology is to explicate the former notions in terms of probabilities, and very often the key idea is to explicate these concepts in terms of probabilistic dependence. For example, according to the standard conceptions of confirmation and coherence, two propositions confirm each other or cohere with each other if and only if they are positively probabilistically dependent. Two propositions disconfirm each other or are incoherent with each other if and only if they are negatively probabilistic dependent (Fitelson 1999, 2001, Shogenji 1999). A second example would be theories of explanatory power. A minimum requirement on the relation between explanandum and explanans is that they are positively probabilistically dependent (Crupi & Tentori 2012, Popper 1959, Schubpach & Sprenger 2011). A third example would be conceptions of unification, according to which a theory unifies two pieces of evidence if it renders them positively probabilistically relevant to each other, i.e. they are probabilistically independent a priori but positively probabilistically dependent in the light of the hypothesis (Myrvold 2003, Schupbach 2005). In a next step Bayesian philosophers of science and epistemologists seek to provide comparative and quantitative notions of confirmation, coherence, explanatory and unificatory power, etc. For this purpose it might be equally useful to study comparative or quantitative notions of probabilistic dependence (i.e. quantitative notions of association, or deviation from independence) as it is useful to study different interpretations of probability.

Measures of probabilistic dependence or deviation from independence have been studied since the early days of the probabilistic approach to understanding rational reasoning. One of the first authors to discuss measures of probabilistic dependence between propositions (instead the dependence of random variables) was John Maynard Keynes. In his 1921 *A Treatise on Probability* Keynes introduces one particular measure of probabilistic dependence that is highly significant for the philosophy of science and epistemology: Keynes's coefficient of dependence.[1] It is highly significant because one can show that Keyne's coefficient of dependence is relevant for understanding almost any Bayesian attempt to capture or explicate interesting aspects of scientific reasoning in terms of probabilities.

Definition 1 (Probabilistic Dependence: Keynes's Coefficient of Dependence)**.** [2]

[1] Keynes (1921) attributes this measure to William Ernest Johnson's manuscript *Cumulative Formula* which was unpublished at that time. The author could not verify whether the latter manuscript has been published since.

[2] Wheeler (2009) calls this the Wayne-Shogenji correlation measure, not because Wayne or Shogenji invented it, but because of the conflicting interpretations that have recently been attached to it by these authors. Wayne (1995) tentatively suggests $\mathfrak{pd}$ as a similarity measure. Shogenji (1999) interprets it as a coherence measure. In the following no such interpretation is presupposed. Since many philosophers before Shogenji and Wheeler used this measure – such as Keynes (1921), Mackie (1969) and Horwich (1982) – I refrain from following Wheeler (2009) in calling it the Wayne-Shogenji correlation measure. In contrast to Wheeler (2009) and Brössel (2013c) I also cautiously refrain from calling it a measure of correlation (I am grateful to an anonymous referee for comments which have made me cautious on this score.). Finally, it is worth mentioning that one can also find alternative measures in the literature. In information theory the most common way to measure the distance between probability distributions is the Kullback-Leibler Divergence. Applied as a measure of deviation from independence, it measures the distance between the actual probability of the conjunction and the independent distribution. In particular, as a measure of deviation from independence the Kullback-Leibler Divergence looks like this:

Definition (Probabilistic Dependence: Kullback-Leibler Divergence)**.**

$$\mathfrak{pd}_{KL}(A_1, \dots, A_n) = \left[\Pr(A_1 \cap \dots \cap A_n) \times \log[\mathfrak{pd}_K(A_1, \dots, A_n)]\right].$$

As one can see, this measure of probabilistic dependence depends crucially on Keynes's coefficient of dependence $\mathfrak{pd}_K$. Thus, studying Keynes's coefficient of dependence is more fundamental. A third measure of deviation from independence is the following:

$$\mathfrak{pd}_K(A_1, \ldots, A_n) = \frac{\Pr(A_1 \cap \ldots \cap A_n)}{\Pr(A_1) \times \ldots \times \Pr(A_n)}$$

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

According to Keynes (1921), the *coefficient of dependence* between propositions $A_1, \ldots, A_n$ gauges the degree of probabilistic dependence of propositions $A_1, \ldots, A_n$, or the degree of deviation from probabilistic independence. One also can define a conditional variant of $\mathfrak{pd}_K$.

Definition 2 (Probabilistic Dependence: Conditional Coefficient of Dependence)**.** [3]

$$\mathfrak{pd}_K(A_1, \ldots, A_n|B) = \frac{\Pr(A_1 \cap \ldots \cap A_n|B)}{\Pr(A_1|B) \times \ldots \times \Pr(A_n|B)}$$

if $\Pr(A_i|B) > 0$ *for all* $1 \leq i \leq n$*, and 0 otherwise.*

The following sections demonstrate that Keynes's coefficient of dependence $\mathfrak{pd}_K$ is intimately related to many central notions within Bayesian epistemology and that it has properties that make it an important tool in the Bayesian epistemologists' tool box for showing that scientific reasoning is best explicated in terms of probability theory.

1.3 Keynes's Coefficient of Dependence and Confirmation

Confirmation theory is one of the central fields of application of the Bayesian machinery and it is intimately related to probabilistic dependence. Bayesian confirmation theory holds that some evidence E confirms a theory T relative to probability function Pr if and only if $\Pr(T|E) > \Pr(T)$. Hence, E confirms T just in case E and T are positively probabilistically dependent relative to probability measure Pr.

According to some proponents of confirmation theory, the relation between probabilistic dependence and confirmation is even stronger: how strongly the theory is confirmed by the evidence depends on how strongly both are probabilistically dependent. This section briefly hints at some confirmation measures that have been suggested in the literature to support this statement. However, none of these measures are defended or rejected (for an overview about this aspect of the discussion see Brössel 2012, 2013a, Crupi 2014, and Fitelson 1999, 2001). For present purposes it is sufficient to make plausible that confirmation and probabilistic dependence in the sense of Keynes's coefficient of dependence are closely related. Proposed confirmation measures that link confirmation very intimately to Keynes's coefficient of dependence are the following:

Definition 3 (Confirmation)**.**

Definition (Deviation From Independence: Difference Measures)**.**

$$\mathfrak{pd}_d(A_1, \ldots, A_n) = \Pr(A_1 \cap \ldots \cap A_n) - \Pr(A_1) \times \ldots \times \Pr(A_n)$$

The difference measure is usually applied when one is interested in the deviation from independence of two propositions. Its generalization for measuring the dependence of two random variables is known as the measure of covariance between the random variables. Unfortunately, discussing and comparing these measures of probabilistic dependence goes beyond the scope of the present paper. This topic will be the content of subsequent research.

[3]Keynes (1921) actually introduces only the conditional variant of $\mathfrak{pd}_K$.

(Carnap 1962)	$d(T,E) = \Pr(T\|E) - \Pr(T)$ *if* $\Pr(E) > 0$
(Mortimer 1988)	$M(T,E) = \Pr(E\|T) - \Pr(E)$ *if* $\Pr(T) > 0$
(Christensen 1999, Joyce 1999)	$S(T,E) = \Pr(T\|E) - \Pr(T\|\overline{E})$ *if* $1 > \Pr(E) > 0$
(Nozick 1981)	$N(T,E) = \Pr(E\|T) - \Pr(E\|\overline{T})$ *if* $1 > \Pr(T) > 0$

(Milne 1996)
$$r(T,E) = \begin{cases} \log\left[\dfrac{\Pr(T|E)}{\Pr(T)}\right] & \text{if } \Pr(T|E) > 0 \\ -\infty & \text{if } \Pr(T|E) = 0 \end{cases}$$

(Crupi et. al. 2007)
$$Z(T,E) = \begin{cases} \frac{d(T,E)}{1-\Pr(T)} & \text{if } \Pr(T|E) \geq \Pr(T) > 0 \\ \frac{d(T,E)}{1-\Pr(\overline{T})} & \text{if } \Pr(T|E) < \Pr(T) \\ 1 & \text{if } \Pr(T) = 0 \end{cases}$$

(Fitelson 2001, Good 1960)
$$l(T,E) = \begin{cases} \log\left[\frac{\Pr(E|T)}{\Pr(E|\overline{T})}\right] & \text{if } \Pr(E|T) > 0 \text{ and } \Pr(E|\overline{T}) > 0 \\ \infty & \text{if } \Pr(E) > 0 \text{ and } \Pr(E|\overline{T}) = 0 \\ -\infty & \text{if } \Pr(E) = 0 \text{ or } \Pr(E|T) = 0 \end{cases}$$

For all these measures it is trivial to substantially link Keynes's coefficient of dependence with confirmation.[4]

[4]The following corollary and all other corollaries and theorems are proved in the appendix.

Corollary 1 (Confirmation)**.**

(Carnap 1962) $$d(T,E) = [\mathfrak{pd}_K(T,E) - 1]\,\Pr(T) \quad \textit{if } \Pr(E) > 0$$

(Mortimer 1988) $$M(T,E) = [\mathfrak{pd}_K(T,E) - 1]\,\Pr(E) \quad \textit{if } \Pr(T) > 0$$

(Christensen 1999, Joyce 1999) $$S(T,E) = \left[\frac{\mathfrak{pd}_K(T,E)-1}{1-\Pr(E)}\right]\Pr(T) \quad \textit{if } 1 > \Pr(E) > 0$$

(Nozick 1981) $$N(T,E) = \left[\frac{\mathfrak{pd}_K(T,E)-1}{1-\Pr(T)}\right]\Pr(E) \quad \textit{if } 1 > \Pr(T) > 0$$

(Milne 1996) $$r(T,E) = \begin{cases} \log[\mathfrak{pd}_K(T,E)] & \text{if } \Pr(T|E) > 0 \\ -\infty & \text{if } \Pr(T|E) = 0 \end{cases}$$

(Crupi et. al. 2007) $$Z(T,E) = \begin{cases} [\mathfrak{pd}_K(T,E) - 1] \times \frac{\Pr(T)}{\Pr(\overline{T})} & \text{if } \Pr(T|E) \geq \Pr(T) > 0 \\ [\mathfrak{pd}_K(T,E) - 1] & \text{if } \Pr(T|E) < \Pr(T) \\ 1 & \text{if } \Pr(T) = 0 \end{cases}$$

(Fitelson 2001, Good 1960) $$l(T,E) = \begin{cases} \log\left[\frac{\mathfrak{pd}_K(T,E)}{\mathfrak{pd}_K(\overline{T},E)}\right] & \text{if } \Pr(E|T) > 0 \text{ and } \Pr(E|\overline{T}) > 0 \\ \infty & \text{if } \Pr(E) > 0 \text{ and } \Pr(E|\overline{T}) = 0 \\ -\infty & \text{if } \Pr(E) = 0 \text{ or } \Pr(E|T) = 0 \end{cases}$$

It is not only easy to demonstrate that the above confirmation measures depend on Keynes's coefficient of dependence between the theory and the evidence. The same holds for all (conceptually) possible confirmation measures. They all depend on Keynes's coefficient of dependence between the theory and the evidence. This is because confirmation depends on the disparity between either the *a priori* probability of the theory and its *a posteriori* probability in the light of the evidence or the *a posteriori* probability of the theory given the evidence and given the negation of the evidence. From these considerations alone it should be clear that it might be very illuminating to study Keynes's coefficient of dependence to reach a better understanding of confirmation. For this to hold true, it is not necessary that confirmation needs to be explicated explicitly by referring to Keynes's coefficient of dependence. Consider the analogous case of coherence and confirmation and suppose that each of these notions can be explicated without explicit reference to the other notion. Nevertheless, studying the relation between coherence and confirmation can provide a better understanding of confirmation and of coherence. The same holds true for measures of confirmation and probabilistic dependence. Even if the quantitative notion of confirmation can be explicated without explicit reference to Keynes's coefficient of dependence, one can nevertheless study the relation between degrees of confirmation and Keynes's coefficient of dependence, and studying this relationship provides a better understanding of confirmation and of Keynes's coefficient of dependence. If this is correct then by studying this relationship (and first, for this purpose, Keynes's coefficient of dependence itself) one can further our understanding of confirmation even if one is not interested in studying Keynes's coefficient of dependence per se. For example, the results mentioned in Corollary 1 help us to understand what confirmation functionally depends on, besides probabilistic dependence in the sense of Keynes's coefficient of dependence. As a simple instance, consider Carnap's measure of confirmation d. According to this measure, if we compare two theories T_1 and T_2 that are to the same degree positively probabilistically dependent with the evidence ($\mathfrak{pd}_K(T_1, E) = \mathfrak{pd}_K(T_2, E) > 1$), then, ceteris paribus, the more plausible theory is the better confirmed one. This verdict is very plausible. For Mortimer's measure one cannot find such a dependence and on this basis one might argue that it is inadequate as a measure of confirmation.

1.4 Keynes's Coefficient of Dependence and Coherence

The study of Keynes's coefficient of dependence is also of importance to Bayesian coherence theory. According to many Bayesian coherentists, two propositions cohere with each other if and only if they are positively probabilistically relevant for each other (Douven & Meijs 2007, Fitelson 2003, Schupbach 2011b, Shogenji 1999).

Shogenji (1999) goes furthest by arguing that coherence is nothing else but Keynes's coefficient for dependence. In particular, Shogenji's (1999) definitions of coherence and of a coherence measure are the following:

Definition 4 (Shogenji Coherence 1)**.**

$$A_1, \ldots, A_n \text{ are coherent if and only if } \mathfrak{pd}_k(A_1, \ldots, A_n) > 1$$

if $\Pr(A_i) > 0$*, and 0 otherwise.*

Definition 5 (Shogenji Coherence 2)**.**

$$Coh_S(A_1, \ldots, A_n) = \mathfrak{pd}_k(A_1, \ldots, A_n)$$

if $\Pr(A_i) > 0$*, and 0 otherwise.*

Other philosophers follow Douven & Meijs (2007) and Fitelson (2003) in defining coherence via confirmation. They hold that the more two propositions mutually confirm each other, the more they cohere. This shows that understanding Keynes's coefficient of dependence is relevant for coherence, since the study of the latter is relevant for the understanding of confirmation. Philosophers like Bovens & Hartmann (2003: 53) and Olsson (2002: 262), who do not define coherence via Keynes's (or any other) coefficient of dependence or confirmation, still admit that the positive probabilistic dependence of two propositions has a positive impact on their coherence. According to them, positive probabilistic relevance in the sense of Keynes's coefficient of dependence increases the coherence of the evidence at least *ceteris paribus.*

1.5 Keynes's Coefficient of Dependence, Explanatory and Unificatory Power, and Diversity of Evidence

Further interesting aspects of scientific reasoning depend on whether and how strongly two propositions are probabilistically dependent in the sense of Keynes's coefficient of dependence. The following paragraphs discuss explanatory and unificatory power, and the diversity of evidence.

Explanatory Power Popper (1959) was one of the first philosophers to suggest a measure for the explanatory power provided by a theory with respect to some evidence. Such a measure should quantify how well a theory explains the evidence. Popper proposes a measure of explanatory power that is ordinally equivalent to the following one by Good (1960):[5]

Definition 6 (Explanatory Power 1)**.**

$$EP_1(T, E) = \frac{\Pr(E|T)}{\Pr(E)}$$

if $\Pr(T) > 0$ *and* $\Pr(E) > 0$*.*

In support of Popper (1959), Good (1960) and McGrew (2003) defend measures of explanatory power that are ordinally equivalent to EP_1. Schupach & Sprenger (2011) propose an alternative measure. They suggest the following measure of the explanatory power of a theory regarding some evidence:

Definition 7 (Explanatory Power 2)**.**

$$EP_2(T, E) = \left[\frac{\Pr(T|E) - \Pr(T|\overline{E})}{\Pr(T|E) + \Pr(T|\overline{E})}\right]$$

if $\Pr(T) > 0$ *and* $1 > \Pr(E) > 0$*.*

[5] The original formulation of Popper's (1959) measure of explanatory power is this: $EP_P(T, E) = \frac{\Pr(E|T) - \Pr(T)}{\Pr(E|T) + \Pr(T)}$.

According to Schupbach, measures of explanatory power are closely related to confirmation measures.[6] In particular, "these measures are structurally equivalent to the confirmation measures" Schupbach (2011a, 814). Measures of confirmation quantify how much the evidence increases the probability of the hypothesis, measures of explanatory power quantify how much the explanans (the hypothesis) increases the probability of the explanandum (the evidence). The measure of explanatory power endorsed by Schupbach & Sprenger (2011), for example, is structurally equivalent to l, the measure of factual support proposed by Kemeny & Oppenheim (1952). Some philosophers argue that l cannot be an adequate measure of confirmation (Milne 1996), while others argue that there is more than one adequate confirmation measure (Joyce 1999). This provokes the question of whether this has implications for theories of explanatory power. Is it possible that some other confirmation measure is the adequate choice for quantifying explanatory power? Is there more than one adequate measure of explanatory power? Popper (1959), Good (1960), McGrew (2003), Schupbach & Sprenger (2011), Schupbach (2011a), and Crupi & Tentori (2012) deal with these questions from various perspectives and provide detailed arguments. For the present purpose it suffices to recognize that the study of probabilistic dependence is often used by philosophers of science to measure explanatory power. For the presented measures of explanatory power EP_1 and EP_2, the exact relation to probabilistic relevance in the sense of Keynes's coefficient of dependence is given by the following corollary:

Corollary 2.

$$EP_1(T,E) = \mathfrak{pd}_k(T,E)$$

if $\Pr(T) > 0$ *and* $\Pr(E) > 0$.

$$EP_2(T,E) = \frac{\mathfrak{pd}_k(E,T) - \mathfrak{pd}_k(\overline{E},T)}{\mathfrak{pd}_k(E,T) + \mathfrak{pd}_k(\overline{E},T)}$$

if $\Pr(T) > 0$ *and* $1 > \Pr(E) > 0$.

Unificatory Power The unificatory power of a theory regarding pieces of evidence is another important concept that has been proposed to be understood in close relation to probabilistic dependence. On Myrvold's (2003: 399) account[7], "the ability of a theory to unify phenomena consists in its ability to render what, on prior grounds, appear to be independent phenomena informationally relevant to each other." Suppose the evidence E consists of two pieces of evidence e_1 and e_2. Myrvold (2003) proposes to measure the unificatory power of a theory T with respect to these pieces of evidence as follows:

Definition 8 (Unificatory Power)**.**

$$UP(e_1,e_2;T) = \log\left[\frac{\frac{\Pr(e_1\cap e_2|T)}{\Pr(e_1|T)\times\Pr(e_2|T)}}{\frac{\Pr(e_1\cap e_2)}{\Pr(e_1)\times\Pr(e_2)}}\right]$$

if $\Pr(T) > 0$,$\Pr(e_1|T) > 0$, *and* $\Pr(e_2|T) > 0$.

Since this measure of unificatory power is complicated, let us set out its relation to Keynes's coefficient of dependence before analyzing it. The relation between UP and $\mathfrak{pd}_k$ is this:

Corollary 3.

$$UP(e_1,e_2;T) = \log\left[\frac{\mathfrak{pd}_k(e_1,e_2|T)}{\mathfrak{pd}_k(e_1,e_2)}\right]$$

if $\Pr(T) > 0$,$\Pr(e_1|T) > 0$, *and* $\Pr(e_2|T) > 0$.

[6] According to Hájek & Joyce (2008), a measure ordinally equivalent to EP_2 can itself be considered a confirmation measure. However, apart from this paper, I am not aware of any other philosophical text which suggests this.

[7] Independently of Myrvold, McGrew (2003) discusses an account of unification (or, as McGrew puts it, of theoretical consilience) that is effectively equivalent to Myrvold's account. For a detailed discussion of both approaches and their connection see Schupbach (2005).

This shows that UP is the ratio of the probabilistic dependence of e_1 and e_2 under the condition T ($\mathfrak{pd}_k(e_1, e_2|T)$) and the probabilistic dependence of e_1 and e_2 unconditionally ($\mathfrak{pd}_k(e_1, e_2)$). Thus, according to Myrvold (2003), the unificatory power of some theory consists in its ability to increase the probabilistic dependence of the pieces of evidence. Increasing the probabilistic dependence of the pieces of evidence in the sense of Keynes's coefficient of dependence is what Myrvold (2003) calls rendering phenomena informationally relevant to each other.

Diversity of Evidence Confirmation theorists often consider it to be uncontroversial that hypotheses are confirmed more by diverse evidence than by uniform evidence. For example, the hypothesis that all ravens are black is much more confirmed if one considers black ravens which were drawn randomly from all over the world than if one considers ravens that were drawn randomly from some small valley in the Black Forest. The reasoning behind this judgement is that it might be the case that only the ravens in that particular valley are all black and the ravens in other valleys are not black. Thus, confirmation theorists need a criterion to distinguish between diverse evidence and uniform or similar evidence. Suppose again that the evidence E consists of two pieces of evidence e_1 and e_2. Since Carnap (1962), philosophers have taken it for granted that if e_1 and e_2 are similar pieces of evidence then they are probabilistically dependent. In the wake of Howson & Urbach (1989), philosophers try to explicate similarity of pieces of evidence according to this property. Along the lines of Howson & Urbach (1989) and Earman (1992), Wayne (1995) tentatively proposes (but does not endorse) the following measure of uniformity or similarity of two pieces of evidence:

Definition 9 (Similarity of Evidence)**.**

$$SE(e_1, e_2) = \frac{\Pr(e_1 \cap e_2)}{\Pr(e_1) \times \Pr(e_2)}$$

if $\Pr(e_1) > 0$ *and* $\Pr(e_2) > 0$.

The relation to Keynes's coefficient of dependence is again trivial and is given by the following corollary:

Corollary 4.

$$SE(e_1, e_2) = \mathfrak{pd}_K(e_1, e_2)$$

if $\Pr(e_1) > 0$ *and* $\Pr(e_2) > 0$.

This shows how the measure SE of the similarity of pieces of evidence can be related to Myrvold's (2003) measure of unificatory power. According to Wayne (1995), the higher the probabilistic dependence of the pieces of evidence in the sense of Keynes's coefficient of dependence, the more similar they are. Hence, an alternative understanding of Myrvold's (2003) measure of the unificatory power of a theory is that it quantifies the degree to which a theory increases the similarity of the pieces of evidence. If the pieces of evidence are more similar given the theory than unconditionally the theory is considered to be unifying the pieces of evidence.

This section set out the relevance of studying Keynes's coefficient of dependence. The hope is that the study of the coefficient of dependence will help us to gain a better understanding of various aspects of scientific reasoning that are central to the philosophy of science and epistemology. Note, it is not presupposed that these notions must be explicitly defined or explicated in terms of Keynes's coefficient of dependence. It also does not imply that whether a given explication is adequate (for example, whether d is an adequate confirmation measure) depends on whether Keynes's coefficient of dependence satisfies certain necessary and sufficient desiderata for measures of probabilistic dependence. All it requires is that these concepts have the following property: if A confirms, coheres with, explains, or is similar to, etc. B, then A and B are positively probabilistically dependent in the sense of Keynes's coefficient of dependence and that, all other things being equal, the more they are positively probabilistically dependent the higher is the degree of confirmation, coherence or explanatory power, etc. If this correct, then studying Keynes's coefficient of dependence is useful. On the one hand Keynes's coefficient of dependence is relevant for confirmation, coherence, explanatory power, etc. and on the other hand it is simpler and clearer than, for example, the conception of confirmation. The reason for this is that the former is only meant to capture the degree of probabilistic dependence. However, in the context of confirmation, coherence, explanatory and unificatory power, etc., other aspects play an important role too. In the context of confirmation, for example, typical assumptions are that confirmed hypotheses are belief-worthy, that old evidence can confirm a hypothesis, that irrelevant conjunctions are less confirmed than the relevant conjunct, etc. Similar points can be made considering concepts such as coherence and explanatory power. For these reasons studying probabilistic dependence in general and Keynes's coefficient of dependence first, is fruitful for Bayesian epistemology.

2 Keynes's Coefficient of Dependence and its Role in Bayesian Epistemology

As a first step, the following Subsection 2.1 investigates whether there is a relation between Keynes's coefficient of dependence and the primary aim of scientific inquiry: finding the truth. In particular, Subsection 2.1 investigates whether the probabilistic dependence of a theory T and the evidence E in the sense of Keynes's coefficient of dependence is an indicator of the truth of T. However, in scientific contexts theories are usually compounds of different hypotheses. Similarly, the evidence is a collection of different pieces of evidence. Hence, a theory T is the conjunction of hypotheses $h_1, \ldots, h_n$, and the evidence E is the conjunction of the pieces of evidence $e_1, \ldots, e_m$. This distinction between the theory and its hypotheses and between the evidence and the pieces of evidence makes the probabilistic dependence in the sense of Keynes's coefficient of dependence especially worth investigating. How does the probabilistic dependence of the pieces of evidence $e_1, \ldots, e_m$ in the sense of Keynes's coefficient of dependence affect the probabilistic dependence of T and E? How does the probabilistic dependence of the hypotheses $h_1, \ldots, h_n$ affect the probabilistic dependence of T and E? In the second and third step, Subsections 2.2 and 2.3 focus on these questions.

2.1 Keynes's Coefficient of Dependence and Truth

One epistemologically interesting feature of Keynes's coefficient of dependence $\mathfrak{pd}_K$ is that the probability of the conjunction of the propositions $A_1, \ldots A_n$ is a monotone function of their prior probabilities and their coefficient of dependence. In addition, the probability of the conjunction of the propositions $A_1, \ldots A_n$ given some proposition B is a monotone function of these propositions' individual conditional probability given B and their conditional coefficient of dependence.

Theorem 1.

$$\Pr(A_1 \cap \ldots \cap A_n) = \mathfrak{pd}_K(A_1, \ldots A_n) \times \prod_{1 \leq i \leq n} \Pr(A_i)$$

$$\Pr(A_1 \cap \ldots \cap A_n | B) = \mathfrak{pd}_K(A_1, \ldots A_n | B) \times \prod_{1 \leq i \leq n} \Pr(A_i | B)$$

It this feature of the coefficient of dependence and its conditional variant that interests Keynes, because "[t]hese coefficients thus belong by definition to a general class of operators, which we may call *separative factors*" (Keynes 1921: 170, emphasis in the original). On the one hand, this property renders the coefficients of dependence useful if one already knows the (conditional) probability of each of the hypotheses $h_1, \ldots, h_n$ and one wants to determine the (conditional) probability of the theory consisting of these hypotheses. On the other hand, and this is more important in the present context, this property renders the coefficient of dependence truth-conducive in the following weak sense: given equal prior or posterior probabilities the truth of all elements of a set with a higher coefficient of dependence is also more likely.

However, if the two true propositions A and B are positively probabilistically dependent, then the two false propositions $\overline{A}$ and $\overline{B}$ are also positively probabilistically dependent.[8] Thus, Keynes's coefficient of dependence *per se* is not an indicator of the truth of propositions. The question is whether there is a stronger relation between truth and Keynes's coefficient of dependence.

In the light of this observation, Shogenji suggests that a high coefficient of dependence (respectively coherence) with true propositions is truth-conducive.

> To put it in a more familiar setting, if a belief is pairwise coherent with a body of beliefs [i.e., if it is positive probabilistic dependent on it] – say, about a fairy – then they have a tendency of being false together, and therefore the belief in question is more likely to be true provided the fairy tale is true; but they also have the tendency of being false together, and therefore the belief in question is more likely to be false if the fairy tale is false. Thus, coherence [positive probabilistic dependence] *with truth* is truth conducive on the level of individual beliefs, but coherence per se is not. (Shogenji 1999: 345)

Still, one could argue that the connection between probabilistic dependence and truth is too weak. In particular, one could point out that Shogenji only argues that probabilistic dependence is probability-conducive, not that it is truth-conducive. After all, a high coefficient of dependence with a true proposition does not guarantee the truth of the proposition in question, only that it has a high probability or a tendency of being true.[9] The question remains: is there a stronger relation between truth and probabilistic dependence in the sense of Keynes's coefficient of dependence?

It is easy to establish a closer link between the truth of a theory and Keynes's coefficient of dependence between that theory and the evidence by referring to convergence theorems of, for example, Gaifman & Snir (1982) or Schervish & Seidenfeld (1990).[10] According to these theorems, the probability of some theory converges to its truth value if the evidence is informative enough to separate the possibilities.[11] By employing these convergence theorems in the study of Keynes's coefficient of dependence between some theory and the evidence the following lemma is provable.[12]

Lemma 1 (Convergence of Keynes's Coefficient of Dependence)**.** Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified in more detail in the appendix on p. 23. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
Then there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories $T \in \mathcal{A}$.

$$\lim_{n\to\infty} \mathfrak{pd}_K(T, E_n^w) = \frac{1}{\Pr(T)} \text{ if } w \vDash T \text{ and } 0 \text{ otherwise.}$$

[8] This has been noted already by Shogenji (1999). Of course, Shogenji is discussing this property in the context of Bayesian coherentism and he takes $\mathfrak{pd}_k$ to be a measure of coherence.

[9] Branden Fitelson argued to this effect in a private e-mail exchange.

[10] Unfortunately, in such a short paper it is not possible to provide a detailed exposition of the mathematically intricate convergence theorems. Accordingly, this paper will presuppose previous acquaintance with these convergence theorems. I refer the interested reader to Hawthorne's (2014) very intelligible exposition of an approach to arrive at one of these convergence results.

[11] A sequence of pieces of evidence separates the set of possibilities W if and only if for every pair of worlds w_i and $w_j \in W$ (with $w_i \neq w_j$) there is one piece of evidence in the sequence such that it is true in one of the possibilities and false in the other.

[12] This lemma and the following theorem have been anticipated and Huber (2005, 2008). See also Brössel (2008, 2013b).

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

This result shows that Keynes's coefficient of dependence between a true theory and the evidence increases the more pieces of evidence one collects. In the long run, Keynes's coefficient of dependence between a true theory T and the evidence E increases to the maximum coefficient of dependence between theory T and evidence E that T can display with respect to some evidence, namely $\frac{1}{Pr(T)}$. Thus, the less plausible a true theory is *a priori*, the higher is the coefficient of dependence between that theory and the evidence in the long run. False theories display the minimal coefficient of dependence in the long run.

This formal result about Keynes's coefficient of dependence in the long run has important ramifications for its application in the philosophy of science and epistemology. In particular, Keynes's coefficient of dependence $\mathfrak{pd}_K$ satisfies requirements on good measures of confirmation that have been suggested by Hempel (1960), Levi (1967), and Huber (2008): (*i*) $\mathfrak{pd}_K$ favors true theories over false theories and (*ii*) $\mathfrak{pd}_K$ favors logically stronger, (i.e. more informative) true theories over logically weaker, (i.e. less informative) true theories after receiving finitely many pieces of evidence and for every additional piece of evidence thereafter.[13]

Theorem 2. Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified in more detail in the appendix on p. 23. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e^w_i = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
Then there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories T_1 and T_2 of $\mathcal{A}$.

1. If $w \vDash T_1$ and $w \vDash \overline{T_2}$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, E^w_m) > \mathfrak{pd}_K(T_2, E^w_m)]$
2. If $w \vDash T_1 \cap T_2$ and $T_1 \vDash T_2$ but $T_2 \nvDash T_1$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, E^w_m) > \mathfrak{pd}_K(T_2, E^w_m)]$

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

Theorem 2 shows that if one compares two theories, one of which is true and the other false, then the coefficient of dependence between the true theory and the evidence is higher than the coefficient of dependence between the false theory and the evidence (after receiving finitely many pieces of evidence and for every piece of evidence thereafter). Thus, Keynes's coefficient of dependence $\mathfrak{pd}_K$ is in a strong sense truth-conducive: it leads us to true theories after receiving finitely many pieces of evidence. It also shows that if one compares two theories, both of which are true but where one of them is logically stronger, then the coefficient of dependence between the logically stronger theory and the evidence is higher than the coefficient of dependence between the logically weaker theory and the evidence (after receiving finitely many pieces of evidence and for every piece of evidence thereafter). This answers the question with respect to the connection between Keynes's coefficient of dependence between a theory and the evidence and the truth of that theory. It also shows that Keynes's coefficient of dependence between a theory and the evidence is more than an indicator of the truth of the theory. It also indicates how informative the true theory is.

However, it is important to note that the above strong results on the truth-conduciveness of Keynes's coefficient of dependence are based on strong assumptions about the availability of evidence, respectively about our information gathering processes. First, in the present context it

[13]Note that Theorem 2 restricts these claims: $\mathfrak{pd}_K$ satisfies both conditions only almost surely: it only holds for every $w \in W'$ for some W' such that $W' \subseteq W$ and $\Pr^*(W') = 1$. It does not necessarily hold for all $w \in W$.

is assumed that our information gathering processes are fully reliable and that we can rely fully on the observational data that we arrive at via these information gathering processes. Although this assumption is fairly standard in the context of theories of confirmation (Brössel 2012, 2013a, Carnap 1962, Fitelson 2001), one must nevertheless emphasize that this assumption is extremely unrealistic. The second and even stronger assumption is that the sequence of possible observational data is informative enough to separate the set W of possibilities, e.g. possible worlds. As already indicated, a sequence of pieces of evidence separates the set of possibilities W if and only if for every pair of possibilities w_i and $w_j \in W$ (with $w_i \neq w_j$) there is one piece of evidence in the sequence such that it is true in one of the possibilities and false in the other. Accordingly, this second assumption requires that for every possibility (in the set of possibilities W) that is not the actual possibility, at least one observation in the sequence of possible observational data reveals to us that it is not the actual possibility.[14] Clearly both assumptions are unrealistic. However, it is also important to note that the unrealistic assumptions make it all the more plausible that a given form of scientific reasoning is rational only if it leads us to true theories, *at least if our information gathering processes are as reliable and informative as described in the above assumptions.* The good news is that Keynes's coefficient of dependence $\mathfrak{pd}_K$ satisfies this minimal requirement.

2.2 Keynes's Coefficient of Dependence Between Pieces of Evidence

Typically the evidence is the conjunction of different pieces of evidence. The question is what impact, if any, combining probabilistically dependent pieces of evidence has on the probabilistic dependence of some theory and the evidence. In outline, the challenge is this: suppose we receive pieces of evidence $e_1, \ldots, e_m$ and we want to know what impact the probabilistic dependence of the pieces of evidence has upon the probabilistic dependence of a theory T and the evidence conceived of as the conjunction of the pieces of evidence. Based on the property that Keynes's coefficient of dependence allows us to separate the impact the (conditional) probabilities of some propositions have on the (conditional) probability of conjunct of these propositions (Theorem 1) this challenge can be answered.

Wayne (1995) presents the following result that displays the relation between the probabilistic dependence of a theory and the evidence in the sense of Keynes's coefficient of dependence on the one hand and the coefficient of dependence of the pieces of evidence and the coefficient of dependence of the theory with each individual piece of evidence on the other.

Theorem 3.

$$\mathfrak{pd}_K(T, e_1 \cap \ldots \cap e_m) = \frac{\Pr(T|e_1)}{\Pr(T)} \times \ldots \times \frac{\Pr(T|e_m)}{\Pr(T)} \times \frac{\frac{\Pr(e_1 \cap \ldots \cap e_m|T)}{\Pr(e_1|T) \times \ldots \times \Pr(e_m|T)}}{\frac{\Pr(e_1 \cap \ldots \cap e_m)}{\Pr(e_1) \times \ldots \times \Pr(e_m)}}$$

$$= \mathfrak{pd}_K(T, e_1) \times \ldots \times \mathfrak{pd}_K(T, e_m) \times \frac{\mathfrak{pd}_K(e_1, \ldots, e_m|T)}{\mathfrak{pd}_K(e_1, \ldots, e_m)}$$

if $\Pr(e_1 \cap \ldots \cap e_m) > 0$.

Myrvold (1996, 2003) and Wheeler (2009) concentrate on the last multiplicand in the theorem. Myrvold (2003) defines his measure UP based on it (see Definition 8). Wheeler (2009), who does not want to make any strong commitment about whether the last multiplicand explicates some important concept of science, dubs this measure 'focused correlation'. Since Wheeler (2009) also assumes that r (i.e. $\log[\mathfrak{pd}_K]$) is an adequate confirmation measure, he concludes that "[t]he focused correlation of e_1 and e_2 relative to a hypothesis T [...] tells us what impact there is on the confirmation of T, if any at all, from combining e_1 and e_2" (Wheeler 2009: 90, notation adapted). However, this is not all the last multiplicand tells us. It also tells us how much a theory increases the coefficient of dependence between the pieces of evidence. This is the reason why Myrvold

[14]For a more detailed discussion of convergence theorems and their philosophical implications especially for scientific realism and anti-realism see Brössel (2013b).

(2003) suggests UP as a measure of the unificatory power of a theory.

Theorem 3 fits our intuitions. First, the coefficient of dependence between a theory and the evidence (i.e. the conjunction of the pieces of evidence) depends on the coefficient of dependence between each individual piece of evidence and the theory. Second, the coefficient of dependence between a theory and the evidence is higher when the theory makes pieces of evidence fit together that did not fit together before – that is, if the coefficient of dependence between pieces of evidence is higher conditional on the theory than on *prior* grounds.[15] It also shows that there is a *ceteris paribus* connection between the coefficient of dependence between the pieces of evidence and the coefficient of dependence between the theory and the evidence. That means that if one holds all other factors in Theorem 3 fixed and only changes the coefficient of dependence between the pieces of evidence, one also changes the coefficient of dependence between the theory and the evidence. The first result is that *ceteris paribus* the coefficient of dependence between the pieces of evidence has a negative effect on the coefficient of dependence between the theory and the evidence.[16] Further, it shows that *ceteris paribus* the coefficient of dependence between the pieces of evidence in light of the theory has a positive effect on the coefficient of dependence between the theory and the evidence. Finally, it shows that *ceteris paribus* the coefficient of dependence between the pieces of evidence has a positive impact on the coefficient of dependence between the theory with each individual piece of evidence.

The preceding paragraphs demonstrate how the coefficient of dependence between the pieces of evidence and the coefficient of dependence between the pieces of evidence given a theory affect the coefficient of dependence between that theory and the evidence and how these considerations might be made fruitful for epistemology and philosophy of science. Now another question is: What is the relation between the coefficient of dependence $\mathfrak{pd}_K(T, e_1, \ldots, e_m)$ between some theory T and the pieces of evidence $e_1, \ldots, e_m$ and the coefficient of dependence between that theory T and the conjunction of the pieces of evidence $e_1 \cap \ldots \cap e_m$? To my knowledge this question has never been addressed in the literature. Note the following corollary which is a variation of a theorem by Shogenji (2007):[17]

[15] Not all readers think that this result fits our intuitions. Accordingly, I at least want to show that it is supported by other intuitions shared by many Bayesian epistemologists. It is supported by our intuitions about unification. More specifically, many Bayesian epistemologists share the intuition that *ceteris paribus* a hypothesis is the more confirmed by the evidence the more it unifies the single pieces of evidence (see section 1.5) and they assume the following: (i) the unificatory power of a theory consists in its ability to increase the coefficient of dependence between the pieces of evidence (McGrew 2003, Myrvold 2003), (ii) the theory is the more confirmed by the evidence the higher the coefficient of dependence between them. Hence, the more the theory in question increases the coefficient of dependence between the pieces of evidence, the higher is the coefficient of dependence between the theory and the conjunction of the pieces of evidence in question.

[16] Again, not all readers will think this result fits our intuitions. Accordingly, I at least want to show that it is supported by other intuitions shared by many Bayesian epistemologists. It is supported by our intuitions about the diversity of evidence. More specifically, many epistemologists share the intuition that *ceteris paribus* hypotheses are confirmed more by diverse pieces of evidence than by uniform pieces of evidence (see section 1.5) and they assume the following: (i) the evidence is the more uniform or similar the more the pieces of evidence are probabilistically dependent (e.g. Myrvold 1995) and (ii) the hypothesis is the more confirmed by the evidence the stronger the probabilistic dependence. Hence, *ceteris paribus* a higher coefficient of dependence between the pieces of evidence has a negative effect on the coefficient of dependence between the theory and the evidence.

[17] In his paper Shogenji (2007) wants to demonstrate why the coherence (in the sense of his coherence measure Coh_S) of mutually independent pieces of evidence *seems to be* truth-conducive for some hypothesis H, even if various impossibility results demonstrated that this cannot be the case (Bovens and Hartmann 2003, Olsson 2002). For this he considers in how far the "degree of coherence between the focal piece (usually the new piece) of evidence and the rest of the evidence is a significant factor in the confirmation of the hypothesis" Shogenji (2007: 367). Accordingly, Shogenji considers the interaction between the coherence/coefficient of dependence of the independent pieces of evidence $e_1, \ldots, e_m$ and the coherence/coefficient of dependence of the independent pieces of evidence $e_1, \ldots, e_{m-1}$ where e_m represents the focal or new piece of evidence and proves that

$$\mathfrak{pd}_K(e_1 \cap \ldots \cap e_{m-1}, e_m) = \frac{\mathfrak{pd}_K(e_1, \ldots, e_m)}{\mathfrak{pd}_K(e_1, \ldots, e_{m-1})}.$$

if $\Pr(e_1 \cap \ldots \cap e_m) > 0$.

Corollary 5.

$$\mathfrak{pd}_K(T, e_1 \cap \ldots \cap e_m) = \frac{\mathfrak{pd}_K(T, e_1, \ldots, e_m)}{\mathfrak{pd}_K(e_1, \ldots, e_m)}$$

if $\Pr(e_1 \cap \ldots \cap e_m) > 0$.

It follows immediately that if the coefficient of dependence between one theory and the pieces of evidence is higher than the coefficient of dependence between another theory and the pieces of evidence then the coefficient of dependence between the first theory and the conjunction of the pieces of evidence is higher than the coefficient of dependence between the latter theory and the conjunction of the pieces of evidence.

Corollary 6.

$$\mathfrak{pd}_K(T_1, e_1 \cap \ldots \cap e_m) > \mathfrak{pd}_K(T_2, e_1 \cap \ldots \cap e_m) \Leftrightarrow$$
$$\mathfrak{pd}_K(T_1, e_1, \ldots, e_m) > \mathfrak{pd}_K(T_2, e_1, \ldots, e_m)$$

Corollary 5 and Theorem 2 imply a further interesting result regarding Keynes's coefficient of dependence. It is truth-conducive even if one considers the coefficient of dependence between a theory and the pieces of evidence and not just the coefficient of dependence between a theory with their conjunction. Since the coefficient of dependence between a theory with the conjunction of the pieces of evidence is truth-conducive (Theorem 2), Corollary 5 shows that the coefficient of dependence between the theory and the pieces of evidence itself is truth-conducive in the strong sense: the coefficient of dependence between a true theory and the pieces of evidence is higher than the coefficient of dependence between a false theory and the pieces of evidence (after receiving finitely many pieces of evidence and for every piece of evidence thereafter). In addition, if two theories are true but one is logically stronger then the coefficient of dependence between the logically stronger one and the pieces of evidence is higher than the coefficient of dependence between the logically weaker one and the pieces of evidence (after receiving finitely many pieces of evidence and for every piece of evidence thereafter).

Theorem 4. Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified in more detail in the appendix on p. 23. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
Then there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories T_1 and T_2 of $\mathcal{A}$.

1. If $w \vDash T_1$ and $w \vDash \overline{T_2}$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, e_1^w, \ldots, e_m^w) > \mathfrak{pd}_K(T_2, e_1^w, \ldots, e_m^w)]$
2. If $w \vDash T_1 \cap T_2$ and $T_1 \vDash T_2$ but $T_2 \nvDash T_1$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, e_1^w, \ldots, e_m^w) > \mathfrak{pd}_K(T_2, e_1^w, \ldots, e_m^w)]$.

The conclusion Shogenji finally reaches is this:

> We have uncovered that when the pieces of evidence are independent with regard to the hypothesis and the rest of the evidence supports the hypothesis, the more coherent the focal piece of evidence is with the rest of the evidence, the more strongly the focal evidence supports the hypothesis. This leaves us with the impression that coherence is truth conducive. (Shogenji 2007: 371)

In the light of Corollary 6 (which does not presuppose that the pieces of evidence are mutually independent) one can see that actually something stronger is true. *Ceteris paribus* the coefficient of dependence between the pieces of evidence has a negative effect on the coefficient of dependence between the theory and the conjunction of the evidence. However, *ceteris paribus* the coefficient of dependence between the pieces of evidence and the theory has a positive effect on the coefficient of dependence between the theory and the conjunction of the evidence.

The proceeding considerations show that the investigation of the Keynes's coefficient of dependence between pieces of evidence and the coefficient of dependence between pieces of evidence and theories is fruitful for the philosophy of science and epistemology. After receiving finitely many pieces of evidence one can identify which theories are true and which false by employing the Keynes's coefficient of dependence $\mathfrak{pd}_K$. Furthermore, Keynes's coefficient of dependence can help to understand what properties diverse or independent pieces of evidence should have and which affect combining such pieces of evidence has on the confirmation of theories. This might help to decide which (series of) experiments a scientist should carry out in order to support or reject theories.

These relations between the coefficient of dependence between the pieces of evidence and the coefficient of dependence between a theory and the evidence are relevant for discussions of the correct explications of confirmation, explanatory and unificatory power, and the diversity of evidence. However, this is not the place to discuss these issues in detail. In this paper the focus is on Keynes's coefficient of dependence and therefore an example may suffice. Consider the confirmation measure l. Corollary 1 shows that the confirmation measure l ($l(T,E) = \frac{\Pr(E|T)}{\Pr(E|\overline{T})}$) can be rewritten solely in terms of Keynes's coefficient of dependence ($l(T,E) = \frac{\mathfrak{pd}_K(T,E)}{\mathfrak{pd}_K(\overline{T},E)}$). In the context of Bayesian updating this quantity is also known as the Bayes factor. Suppose now that the evidence E is the conjunction of the pieces of evidence $e_1, \ldots, e_m$. For the Bayes factor one can show how the Bayes factor of each piece of evidence e_i influences the Bayes factor $l(T,E) = \frac{\mathfrak{pd}_K(T,e_1\cap\ldots\cap e_m)}{\mathfrak{pd}_K(\overline{T},e_1\cap\ldots\cap e_m)}$.

Corollary 7.

$$l(T, e_1 \cap \ldots \cap e_m) = \frac{\mathfrak{pd}_K(T,E)}{\mathfrak{pd}_K(\overline{T},E)} = \frac{\mathfrak{pd}_K(T,e_1) \times \ldots \times \mathfrak{pd}_K(T,e_m) \times \frac{\mathfrak{pd}_K(e_1,\ldots,e_m|T)}{\mathfrak{pd}_K(e_1,\ldots,e_m)}}{\mathfrak{pd}_K(\overline{T},e_1) \times \ldots \times \mathfrak{pd}_K(\overline{T},e_m) \times \frac{\mathfrak{pd}_K(e_1,\ldots,e_m|\overline{T})}{\mathfrak{pd}_K(e_1,\ldots,e_m)}}$$

$$l(T, e_1 \cap \ldots \cap e_m) = l(T,e_1) \times \ldots \times l(T,e_m) \times \frac{\mathfrak{pd}_K(e_1,\ldots,e_m|T)}{\mathfrak{pd}_K(e_1,\ldots,e_m|\overline{T})}$$

This corollary is an example of how Keynes's coefficient of dependence might help us to gain a better understanding of confirmation. In addition, this might be fruitful for the discussion of possible measures of unification. In particular, Bayesians who think that r is an inadequate measure of confirmation and that l is a better measure of confirmation (Brössel 2012, Fitelson 2001, Hawthorne 2014) might argue on the basis of Corollary 7 that unification should better be explicated as follows:

Definition 10 (Unificatory Power)**.**

$$UP_2(e_1,\ldots,e_m;T) = \frac{\mathfrak{pd}_K(e_1,\ldots,e_m|T)}{\mathfrak{pd}_K(e_1,\ldots,e_m|\overline{T})}$$

2.3 Keynes's Coefficient of Dependence Between Hypotheses

As already noted, theories often 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, ..., h_n, similar questions arise regarding the effect the coefficient of dependence between these hypotheses has upon the coefficient of dependence between their conjunction (i.e. the theory) and the evidence. A theorem similar to Theorem 3 is provable for hypotheses:

Theorem 5.

$$\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E) = \frac{\Pr(h_1|E)}{\Pr(h_1)} \times \ldots \times \frac{\Pr(h_n|E)}{\Pr(h_n)} \times \frac{\frac{\Pr(h_1\cap\ldots\cap h_n|E)}{\Pr(h_1|E)\times\ldots\times\Pr(h_n|E)}}{\frac{\Pr(h_1\cap\ldots\cap h_n)}{\Pr(h_1)\times\ldots\times\Pr(h_n)}}$$

$$= \mathfrak{pd}_K(h_1, E) \times \ldots \times \mathfrak{pd}_K(h_n, E) \times \frac{\mathfrak{pd}_K(h_1, \ldots, h_n|E)}{\mathfrak{pd}_K(h_1, \ldots, h_n)}$$

if $\Pr(h_1 \cap \ldots \cap h_n) > 0$.

This result is in accordance with our intuitions. In order to achieve a high coefficient of dependence between the conjunction of the hypotheses and the evidence, the conjunction of the hypotheses should be initially improbable, but probable given the evidence. Consequently, the coefficient of dependence between the hypotheses should be initially low so that their conjunction is *a priori* improbable. However, the coefficient of dependence between the hypotheses given the evidence should be high so that the conjunction is probable given the evidence. Furthermore, the coefficient of dependence between each individual hypothesis and the evidence should be high as well.

Another question is: How does the coefficient of dependence between the hypotheses and the evidence affect the coefficient of dependence between the conjunction of the hypotheses and the evidence? More formally, what is the relation between $\mathfrak{pd}_K(h_1, \ldots, h_n, E)$ and $\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E)$? The following theorem answers this question:

Corollary 8.

$$\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E) = \frac{\mathfrak{pd}_K(h_1, \ldots, h_n, E)}{\mathfrak{pd}_K(h_1, \ldots, h_n)}$$

if $\Pr(h_1 \cap \ldots \cap h_n) > 0$.

This theorem depicts a strong connection between the coefficient of dependence between the hypotheses and the evidence and the coefficient of dependence between the conjunction of the hypotheses and the evidence. The coefficient of dependence between the conjunction of hypotheses and the evidence is nothing else than the ratio of the coefficient of dependence between the hypotheses and the evidence and the coefficient of dependence between the hypotheses.

A theorem that generalizes and summarizes the findings of Corollary 5 and Corollary 8 is the following:

Theorem 6.

$$\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, e_1 \cap \ldots \cap e_m) = \frac{\mathfrak{pd}_K(h_1, \ldots, h_n, e_1, \ldots, e_m)}{\mathfrak{pd}_K(h_1, \ldots, h_n) \times \mathfrak{pd}_K(e_1, \ldots, e_m)}$$

if $\Pr(h_1 \cap \ldots \cap h_n) > 0$ and $\Pr(e_1 \cap \ldots \cap e_m) > 0$.

This theorem shows that the coefficient of dependence between the theory and the evidence depends positively on the coefficient of dependence between the hypotheses and the pieces of evidence. It depends negatively on the coefficient of dependence between the hypotheses and the coefficient of dependence between the pieces of evidence.

From Theorem 2 and Corollary 8 one can derive another important result. The ratio of the coefficient of dependence between the hypotheses and the evidence and the coefficient of dependence between the hypotheses is truth-conducive in the strong sense. In particular, after receiving finitely many pieces of evidence the following holds for Keynes's coefficient of dependence $\mathfrak{pd}_K$: the disparity between the coefficient of dependence between the hypotheses and evidence and the coefficient of dependence between the hypotheses is an indicator of the truth of the hypotheses. The ratio of the coefficient of dependence between hypotheses and the evidence and the coefficient of dependence between the hypotheses is higher for compounds of true hypotheses than for compounds of hypotheses that contain at least one false hypothesis (after finitely many pieces of evidence and for every piece of evidence thereafter). Furthermore, the ratio of the coefficient of dependence between the hypotheses and the evidence and the coefficient of dependence between the hypotheses is higher for logically stronger compounds of true hypotheses than for logically weaker compounds of true hypotheses (after finitely many pieces of evidence and for every piece of evidence thereafter). The following theorem states this conjecture more precisely:

Theorem 7. Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified in more detail in the appendix on p. 23. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \dots, e_n, \dots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let Pr^* be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\mathrm{Pr}^*(A) = \mathrm{Pr}(A)$ for all $A \in \mathcal{A}$.
Then there is a $W' \subseteq W$ with $\mathrm{Pr}^*(W') = 1$ so that the following holds for every $w \in W'$ and all hypotheses $h_1, \dots h_n$ and $h'_1, \dots h'_m$ of $\mathcal{A}$.

1. If $w \vDash h_1 \cap \dots \cap h_n$ and $w \vDash (\overline{h'_1 \cap \dots \cap h'_m})$, then:
$$\exists k \forall l \geq k : \left[\frac{\mathfrak{pd}_K(h_1, \dots, h_n, E_l^w)}{\mathfrak{pd}_K(h_1, \dots, h_n)} > \frac{\mathfrak{pd}_K(h'_1, \dots, h'_m, E_l^w)}{\mathfrak{pd}_K(h'_1, \dots, h'_n)} \right].$$
2. If $w \vDash h_1 \cap \dots \cap h_n \cap h'_1 \cap \dots \cap h'_m$ and $h_1 \cap \dots \cap h_n \vDash h'_1 \cap \dots \cap h'_m$ but $h'_1 \cap \dots \cap h'_m \nvDash h_1 \cap \dots \cap h_n$, then:
$$\exists k \forall l \geq k : \left[\frac{\mathfrak{pd}_K(h_1, \dots, h_n, E_l^w)}{\mathfrak{pd}_K(h_1, \dots, h_n)} > \frac{\mathfrak{pd}_K(h'_1, \dots, h'_m, E_l^w)}{\mathfrak{pd}_K(h'_1, \dots, h'_n)} \right]$$

where $E_l^w = \bigcap_{0 \leq i \leq l} e_i^w$.

These considerations show that studying the coefficient of dependence between the hypotheses of a theory and the evidence is extremely useful for the philosophy of science and epistemology. By studying the coefficient of dependence between the hypotheses and the evidence one can identify which (compounds of) hypotheses are true and which false after receiving finitely many pieces of evidence. In addition, the coefficient of dependence is essential for understanding the inter-theoretic relations which can help a theory to gain support from the evidence. According to some Bayesians the importance of Keynes's coefficient of dependence is even higher. If one follows Popper (1959) and others in relating explanatory power to Keynes's coefficient of dependence, these results are of greatest importance for understanding how theories explain the evidence. If one follows Myrvold (2003) in tying unificatory power to Keynes's coefficient of dependence these results are of greatest importance for understanding how theories unify the evidence.

3 Conclusions

Section 1 displays various connections between Keynes's coefficient of dependence $\mathfrak{pd}_K$ and various essential concepts of the philosophy of science and epistemology. In particular it shows that Keynes's coefficient of dependence is closely related to various proposed Bayesian measures of confirmation, coherence, explanatory and unificatory power, and the diversity of evidence. This renders Keynes's coefficient of dependence one of the most interesting and central factors in the philosophy of science and epistemology. It is imperative that it be investigated further.

The investigation of the formal properties of Keynes's coefficient of dependence $\mathfrak{pd}_K$ in section 2 vindicates the claim that it is itself a useful tool within the philosophy of science and epistemology.

First, Theorem 1 shows that, *ceteris paribus*, Keynes's coefficient of dependence is truth-conducive in the following weak sense: if one holds the prior (respectively the posterior) probabilities of some set of propositions fixed, the prior (respectively the posterior) probability of the truth of all propositions increases with their (conditional) coefficient of dependence.

Second, Theorem 2 demonstrates that the coefficient of dependence between true theories and the evidence is higher than the coefficient of dependence between false theories and the evidence (after the receipt of finitely many pieces of evidence and for every piece of evidence thereafter). This establishes that the coefficient of dependence $\mathfrak{pd}_K$ is truth-conducive in a strong sense. Additionally, it holds that the coefficient of dependence between logically stronger true theories and the evidence is higher than the coefficient of dependence between logically weaker true theories and the evidence (after receiving finitely many pieces of evidence and for every piece of evidence thereafter).

Third, relying on the distinction between the coefficient of dependence and the conditional coefficient of dependence, Theorem 3 shows that the coefficient of dependence between the pieces of evidence has a negative impact on the coefficient of dependence between a theory and the evidence. However, the conditional coefficient of dependence between the pieces of evidence given the theory contributes positively to the coefficient of dependence between the theory and the evidence.

Fourth, Theorem 5 demonstrates that the coefficient of dependence between the hypotheses of a theory does not have a positive impact on the coefficient of dependence between the theory and the evidence; it has a negative impact. The conditional coefficient of dependence between the hypotheses given the evidence contributes positively to the coefficient of dependence between the theory and the evidence.

Fifth, the coefficient of dependence between the hypotheses and the pieces of evidence contributes to the coefficient of dependence between the theory and the evidence. However, the coefficient of dependence between the hypotheses and the coefficient of dependence between the pieces of evidence does not.

Keynes's coefficient of dependence $\mathfrak{pd}_K$ is indeed a key concept of the philosophy of science and epistemology. It is intimately related to essential aspects of scientific reasoning such as confirmation, coherence, explanatory and unificatory power, and the diversity of evidence. Thus, it is imperative to investigate the consequences these results have for the inter-theoretic relations between hypotheses and pieces of evidence. Which exact epistemological consequences these findings might have depends on the exact theories of confirmation, coherence, etc., that one adopts. However, the above theorems demonstrate how important it is to study how Keynes's coefficient of dependence is related to different forms of scientific reasoning. This holds in particular because the coefficient of dependence $\mathfrak{pd}_K$ can be used to discern true from false theories.

We can conclude that Keynes's coefficient of dependence is an important tool in the Bayesian epistemologists' tool box for showing that scientific reasoning can be explicated in terms of probability theory and that one can formulate and justify epistemic norms and evaluations by relying on these explications.

A Proofs of Corollaries

Proof of Corollary 1

1. According to Definition 3:

$$\begin{aligned} d(T,E) &= \Pr(T|E) - \Pr(T) \\ &= \left[\frac{\Pr(T|E)}{\Pr(T)} - 1\right] \times \Pr(T) \\ &= [\mathfrak{pd}_K(T,E) - 1] \times \Pr(T) \end{aligned}$$

2. According to Definition 3:

$$\begin{aligned}
M(T,E) &= \Pr(E|T) - \Pr(E) \\
&= \left[\frac{\Pr(E|T)}{\Pr(E)} - 1\right] \times \Pr(E) \\
&= [\mathfrak{pd}_K(T,E) - 1] \times \Pr(E)
\end{aligned}$$

3. According to Definition 3:

$$\begin{aligned}
S(T,E) &= \Pr(T|E) - \Pr(T|\bar{(}E)) \\
&= \frac{\Pr(T|E) - \Pr(T)}{\Pr(\overline{E})} \\
&= \left[\frac{\frac{\Pr(T|E)}{\Pr(T)} - 1}{\Pr(\overline{E})}\right] \times \Pr(T) \\
&= \left[\frac{\mathfrak{pd}_K(T,E) - 1}{1 - \Pr(E)}\right] \times \Pr(T)
\end{aligned}$$

4. According to Definition 3:

$$\begin{aligned}
M(T,E) &= \Pr(E|T) - \Pr(E|\bar{(}T)) \\
&= \frac{\Pr(E|T) - \Pr(E)}{\Pr(\overline{T})} \\
&= \left[\frac{\frac{\Pr(E|T)}{\Pr(E)} - 1}{\Pr(\overline{T})}\right] \times \Pr(E) \\
&= \left[\frac{\mathfrak{pd}_K(T,E) - 1}{1 - \Pr(T)}\right] \times \Pr(E)
\end{aligned}$$

5. According to Definition 3:

$$r(T,E) = \log\left[\frac{\Pr(T|E)}{\Pr(T)}\right],$$

if $\Pr(T) > 0$ and $\Pr(E) > 0$. Since $\frac{\Pr(T|E)}{\Pr(T)} = \frac{\Pr(T \cap E)}{\Pr(T) \times \Pr(E)}$ it follows with Definition 1 that:

$$r(T,E) = \log\left[\mathfrak{pd}_K(T,E)\right],$$

if $1 > \Pr(T) > 0$ and $\Pr(E) > 0$.

6. According to Definition 3:

$$\begin{aligned}
Z(T,E) &= \begin{cases} \frac{d(T,E)}{1-\Pr(T)} & \text{if } \Pr(T|E) \geq \Pr(T) > 0 \\ \frac{d(T,E)}{1-\Pr(\overline{T})} & \text{if } \Pr(T|E) < \Pr(T) \\ 1 & \text{if } \Pr(T) = 0 \end{cases} \\
&= \begin{cases} \frac{[\mathfrak{pd}_K(T,E)-1] \times \Pr(T)}{1-\Pr(T)} & \text{if } \Pr(T|E) \geq \Pr(T) > 0 \\ \frac{[\mathfrak{pd}_K(T,E)-1] \times \Pr(T)}{\Pr(T)} & \text{if } \Pr(T|E) < \Pr(T) \\ 1 & \text{if } \Pr(T) = 0 \end{cases} \\
&= \begin{cases} [\mathfrak{pd}_K(T,E) - 1] \times \frac{\Pr(T)}{1-\Pr(T)} & \text{if } \Pr(T|E) \geq \Pr(T) > 0 \\ [\mathfrak{pd}_K(T,E) - 1] \times \frac{\Pr(T)}{\Pr(T)} & \text{if } \Pr(T|E) < \Pr(T) \\ 1 & \text{if } \Pr(T) = 0 \end{cases}
\end{aligned}$$

7. According to Definition 3:

$$l(T, E) = \log \left[\frac{\Pr(E|T)}{\Pr(E|\overline{T})} \right],$$

if $1 > \Pr(T) > 0$ and $\Pr(E) > 0$. It holds however that:

$$\log \left[\frac{\Pr(E|T)}{\Pr(E|\overline{T})} \right] = \log \left[\frac{\frac{\Pr(E|T)}{\Pr(E)}}{\frac{\Pr(E|\overline{T}}{\Pr(E)})} \right].$$

This implies

$$l(T, E) = \log \left[\frac{\mathfrak{pd}_K(T, E)}{\mathfrak{pd}_K(\overline{T}, E)} \right],$$

if $1 > \Pr(T) > 0$ and $\Pr(E) > 0$.

Proof of Corollary 2

1. According to Definition 6:

$$\begin{aligned} EP_1(T, E) &= \frac{\Pr(E|T)}{\Pr(E)} \\ &= \frac{\Pr(E \cap T)}{\Pr(E) \times \Pr(T)} \\ &= \mathfrak{pd}_K(T, E) \text{ (with to Definition 1)} \end{aligned}$$

2. According to Definition 7:

$$\begin{aligned} EP_2(T, E) &= \left[\frac{\Pr(T|E) - \Pr(T|\overline{E})}{\Pr(T|E) + \Pr(T|\overline{E})} \right] \\ &= \left[\frac{\frac{\Pr(T|E) - \Pr(T|\overline{E})}{\Pr(T)}}{\frac{\Pr(T|E) + \Pr(T|\overline{E})}{\Pr(T)}} \right] \\ &= \left[\frac{\frac{\Pr(T|E)}{\Pr(T)} - \frac{\Pr(T|\overline{E})}{\Pr(T)}}{\frac{\Pr(T|E)}{\Pr(T)} + \frac{\Pr(T|\overline{E})}{\Pr(T)}} \right] \\ &= \frac{\mathfrak{pd}_K(E, T) - \mathfrak{pd}_K(\overline{E}, T)}{\mathfrak{pd}_K(E, T) + \mathfrak{pd}_K(\overline{E}, T)} \end{aligned}$$

Proof of Corollary 3 According to Definition 8:

$$\begin{aligned} UP(e_1, e_2; T) &= \log \left[\frac{\frac{\Pr(e_1 \cap e_2|T)}{\Pr(e_1|T) \times \Pr(e_2|T)}}{\frac{\Pr(e_1 \cap e_2)}{\Pr(e_1) \times \Pr(e_2)}} \right] \\ &= \log \left[\frac{\mathfrak{pd}_K(e_1, e_2|T)}{\mathfrak{pd}_K(e_1, e_2)} \right] \end{aligned}$$

Proof of Corollary 4 According to Definition 9:

$$\begin{aligned} SE(e_1, e_2) &= \frac{\Pr(e_1 \cap e_2)}{\Pr(e_1) \times \Pr(e_2)} \\ &= \mathfrak{pd}_K(e_1, e_2) \end{aligned}$$

Proof of Corollary 5

$$\mathfrak{pd}_K(T, e_1 \cap \ldots \cap e_m) = \frac{\Pr(T \cap e_1 \cap \ldots \cap e_m)}{\Pr(T) \times \Pr(e_1 \cap \ldots \cap e_m)}$$
$$= \frac{\frac{\Pr(T \cap e_1 \cap \ldots \cap e_m)}{\Pr(T) \times \Pr(e_1) \times \ldots \times \Pr(e_m)}}{\frac{\Pr(e_1 \cap \ldots \cap e_m)}{\Pr(e_1) \times \ldots \times \Pr(e_m)}}$$
$$= \frac{\mathfrak{pd}_K(T, e_1, \ldots, e_m)}{\mathfrak{pd}_K(e_1, \ldots, e_m)}$$

Proof of Corollary 6

$$\mathfrak{pd}_K(T_1, e_1 \cap \ldots \cap e_m) > \mathfrak{pd}_K(T_2, e_1 \cap \ldots \cap e_m) \Leftrightarrow$$
$$\frac{\mathfrak{pd}_K(T_1, e_1, \ldots, e_m)}{\mathfrak{pd}_K(e_1, \ldots, e_m)} > \frac{\mathfrak{pd}_K(T_2, e_1, \ldots, e_m)}{\mathfrak{pd}_K(e_1, \ldots, e_m)} \Leftrightarrow$$
$$\mathfrak{pd}_K(T_1, e_1, \ldots, e_m) > \mathfrak{pd}_K(T_2, e_1, \ldots, e_m).$$

Proof of Corollary 7 Corollary 7 follows trivially from Corollary 1 for the l confirmation measure and Theorem 3.

Proof of Corollary 8

$$\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E) = \frac{\Pr(h_1 \cap \ldots \cap h_n \cap E)}{\Pr(h_1 \cap \ldots \cap h_n) \times \Pr(E)}$$
$$= \frac{\frac{\Pr(h_1 \cap \ldots \cap h_n \cap E)}{\Pr(h_1) \times \ldots \times \Pr(h_n) \times \Pr(E)}}{\frac{\Pr(h_1 \cap \ldots \cap h_n)}{\Pr(h_1) \times \ldots \times \Pr(h_n)}}$$
$$= \frac{\mathfrak{pd}_K(h_1, \ldots, h_n, E)}{\mathfrak{pd}_K(h_1, \ldots, h_n)}$$

B Proofs of Theorems

Proof of Theorem 1.

1.

$$\Pr(A_1 \cap \ldots \cap A_n) = \Pr(A_1 \cap \ldots \cap A_n)$$
$$\Pr(A_1 \cap \ldots \cap A_n) = \frac{\Pr(A_1 \cap \ldots \cap A_n)}{\prod_{1 \le i \le n} \Pr(A_i)} \times \prod_{1 \le i \le n} \Pr(A_i)$$
$$\Pr(A_1 \cap \ldots \cap A_n) = \mathfrak{pd}_K(A_1, \ldots A_n) \times \prod_{1 \le i \le n} \Pr(A_i)$$

2.

$$\Pr(A_1 \cap \ldots \cap A_n | B) = \Pr(A_1 \cap \ldots \cap A_n | B)$$
$$\Pr(A_1 \cap \ldots \cap A_n | B) = \frac{\Pr(A_1 \cap \ldots \cap A_n | B)}{\prod_{1 \le i \le n} \Pr(A_i | B)} \times \prod_{1 \le i \le n} \Pr(A_i | B)$$
$$\Pr(A_1 \cap \ldots \cap A_n | B) = \mathfrak{pd}_K(A_1, \ldots A_n | B) \times \prod_{1 \le i \le n} \Pr(A_i | B)$$

Proof of Theorem 2 The proof proceeds as follows: First, the Gaifman-Snir Theorem is presented (for a proof see Gaifman-Snir 1982). Second, Lemma 1 is proven. Third, Theorem 2 is completed.

1.) The Gaifman-Snir Theorem: Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some language $\mathcal{L}$ suitable for arithmetic. In particular, let $\mathcal{L}$ be some first order language containing the numerals '1', '2', '3', ... as names, respectively, individual constants, and symbols for addition, multiplication, identity etc. In addition, let $\mathcal{L}$ contain finitely many relations and functional symbols. Gaifman and Snir (1982) call them the 'empirical symbols'. Accordingly one can think of the possibilities in W as models for that language $\mathcal{L}$ (which agree on the interpretation of the mathematical symbols but can disagree on the interpretation of the empirical symbols).

Now let e_1,..., e_n, ... be a sequence of propositions of $\mathcal{A}$ that separates W, and for all $w \in W$ let $e_i^w = e_i$, if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
Then there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories T of $\mathcal{A}$:

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

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

2.) Lemma 1: Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified in more detail above. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
Then according to the Gaifman-Snir Theorem there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories T of $\mathcal{A}$:

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

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

Now it holds that

$$\begin{aligned}
\lim_{n\to\infty} \mathfrak{pd}_K(T, E_n^w) &= \lim_{n\to\infty} \frac{\Pr(T|E_n^w)}{\Pr(T)} \\
&= \lim_{n\to\infty} \Pr(T|E_n^w) \times \lim_{n\to\infty} \frac{1}{\Pr(T)} \\
&= \lim_{n\to\infty} \Pr(T|E_n^w) \times \frac{1}{\Pr(T)} \\
&= \begin{cases} \frac{1}{\Pr(T)}, & \text{if } \lim_{n\to\infty} \Pr(T|E_n^w) = 1 \\ 0, & \text{if } \lim_{n\to\infty} \Pr(T|E_n^w) = 0 \end{cases} \\
&= \begin{cases} \frac{1}{\Pr(T)}, & \text{if } w \vDash T_1 \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

3.) Proof of Theorem 2: Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified

in more detail above. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
Then according to Lemma 1 there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories T of $\mathcal{A}$:

$$\lim_{n\to\infty} \mathfrak{pd}_K(T, E_n^w) = \frac{1}{\Pr(T)} \text{ if } w \vDash T \text{ and 0 otherwise.}$$

1.) Let T_1 and T_2 be two theories of $\mathcal{A}$ and suppose additionally that $w \vDash T_1$ and $w \vDash \overline{T_2}$.

We know that $\lim_{n\to\infty} \mathfrak{pd}_K(T_1, E_n^w) = \frac{1}{\Pr(T_1)}$ since $w \vDash T_1$. We also know that $\lim_{n\to\infty} \mathfrak{pd}_K(T_2, E_n^w) = 0$ since $w \vDash \overline{T_2}$.

Let $\epsilon = \frac{\frac{1}{\Pr(T_1)}}{2}$. By the definition of lim it holds that: $\exists n \in \mathbb{N} \forall m \geq n : |\frac{1}{\Pr(T_1)} - \mathfrak{pd}_K(T_1, E_m^{w'})| < \epsilon$ and $\exists n' \in \mathbb{N} \forall m \geq n' |0 - \mathfrak{pd}_K(T_2, E_m^{w'})| < \epsilon$.

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

$$\mathfrak{pd}_K(T_1, E_m^{w'}) > \mathfrak{pd}_K(T_2, E_m^{w'})$$

2.) Now assume that $w \vDash T_1$ and $w \vDash T_2$ and $T_1 \vDash T_2$ but $T_2 \nvDash T_1$.
Because of Lemma 1 we know that $\lim_{n\to\infty} \mathfrak{pd}_K(T_1, E_n^w) = \frac{1}{\Pr(T_1)}$ since $w \vDash T_1$. We also know that $\lim_{n\to\infty} \mathfrak{pd}_K(T_2, E_n^w) = \frac{1}{\Pr(T_2)}$ since $w \vDash T_2$.
The assumption is that Pr is a strict probability function and $T_1 \vDash T_2$ but $T_2 \nvDash T_1$. It follows that $\Pr(T_1) < \Pr(T_2)$ 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: $\exists n \in \mathbb{N} \forall m \geq n : |\frac{1}{p(T_1)} - \mathfrak{pd}_K(T_1, E_m^w)| < \epsilon$ and $\exists n' \in \mathbb{N} \forall m \geq n' |\frac{1}{p(T_2)} - \mathfrak{pd}_K(T_2, E_m^w)| < \epsilon$.

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

$$\mathfrak{pd}_K(T_1, E_m^w) > \mathfrak{pd}_K(T_2, E_m^w)$$

Proof of Theorem 3

$$\mathfrak{pd}_K(T, e_1 \cap \ldots \cap e_m) = \frac{\Pr(e_1 \cap \ldots \cap e_m \cap T)}{\Pr(e_1 \cap \ldots \cap e_m) \times \Pr(T)}$$

$$= \frac{\Pr(T|e_1)}{\Pr(T)} \times \ldots \times \frac{\Pr(T|e_m)}{\Pr(T)} \times \frac{\frac{\Pr(e_1 \cap \ldots \cap e_m|T)}{\Pr(e_1 \cap \ldots \cap e_m)}}{\frac{\Pr(T|e_1)}{\Pr(T)} \times \ldots \times \frac{\Pr(T|e_m)}{\Pr(T)}}$$

$$= \frac{\Pr(T|e_1)}{\Pr(T)} \times \ldots \times \frac{\Pr(T|e_m)}{\Pr(T)} \times \frac{\frac{\Pr(e_1 \cap \ldots \cap e_m|T)}{\Pr(e_1 \cap \ldots \cap e_m)}}{\frac{\Pr(e_1|T) \times \ldots \times \Pr(e_m|T)}{\Pr(e_1) \times \ldots \times \Pr(e_m)}}$$

$$= \frac{\Pr(T|e_1)}{\Pr(T)} \times \ldots \times \frac{\Pr(T|e_m)}{\Pr(T)} \times \frac{\frac{\Pr(e_1 \cap \ldots \cap e_m|T)}{\Pr(e_1|T) \times \ldots \times \Pr(e_m|T)}}{\frac{\Pr(e_1 \cap \ldots \cap e_m)}{\Pr(e_1) \times \ldots \times \Pr(e_m)}}$$

$$= \mathfrak{pd}_K(T, e_1) \times \ldots \times \mathfrak{pd}_K(T, e_m) \times \frac{\mathfrak{pd}_K(e_1, \ldots, e_m|T)}{\mathfrak{pd}_K(e_1, \ldots, e_m)}$$

Proof of Theorem 4 Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified in more detail in the appendix in connection with Theorem 2. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
Then according to Theorem 2 there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories T_1 and T_2 of $\mathcal{A}$.

1. If $w \vDash T_1$ and $w \vDash \overline{T_2}$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, \bigcap_{0 \leq i \leq m} e_i^w) > \mathfrak{pd}_K(T_2, \bigcap_{0 \leq i \leq m} e_i^w)]$
2. If $w \vDash T_1 \cap T_2$ and $T_1 \vDash T_2$ but $T_2 \nvDash T_1$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, \bigcap_{0 \leq i \leq m} e_i^w) > \mathfrak{pd}_K(T_2, \bigcap_{0 \leq i \leq m} e_i^w)]$.

According to Corollary 6 it holds that

$$\mathfrak{pd}_K(T_1, \bigcap_{0 \leq i \leq m} e_i^w) > \mathfrak{pd}_K(T_2, \bigcap_{0 \leq i \leq m} e_i^w) \Leftrightarrow$$
$$\mathfrak{pd}_K(T_1, e_1^w, \ldots, e_m^w) > \mathfrak{pd}_K(T_2, e_1^w, \ldots, e_m^w).$$

Hence, there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all theories T_1 and T_2 of $\mathcal{A}$.

1. If $w \vDash T_1$ and $w \vDash \overline{T_2}$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, e_1^w, \ldots, e_m^w) > \mathfrak{pd}_K(T_2, e_1^w, \ldots, e_m^w)]$
2. If $w \vDash T_1 \cap T_2$ and $T_1 \vDash T_2$ but $T_2 \nvDash T_1$, then:
   $\exists n \forall m \geq n : [\mathfrak{pd}_K(T_1, e_1^w, \ldots, e_m^w) > \mathfrak{pd}_K(T_2, e_1^w, \ldots, e_m^w)]$.

Proof of Theorem 5

$$\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E) = \frac{\Pr(h_1 \cap \ldots \cap h_n \cap E)}{\Pr(h_1 \cap \ldots \cap h_n) \times \Pr(E)}$$

$$= \frac{\Pr(h_1|E)}{\Pr(h_1)} \times \ldots \times \frac{\Pr(h_n|E)}{\Pr(h_n)} \times \frac{\frac{\Pr(h_1 \cap \ldots \cap h_n|E)}{\Pr(h_1 \cap \ldots \cap h_n)}}{\frac{\Pr(h_1|E)}{\Pr(h_1)} \times \ldots \times \frac{\Pr(h_n|E)}{\Pr(h_n)}}$$
$$= \frac{\Pr(h_1|E)}{\Pr(h_1)} \times \ldots \times \frac{\Pr(h_n|E)}{\Pr(h_n)} \times \frac{\frac{\Pr(h_1 \cap \ldots \cap h_n|E)}{\Pr(h_1 \cap \ldots \cap h_n)}}{\frac{\Pr(h_1|E) \times \ldots \times \Pr(h_n|E)}{\Pr(h_1) \times \ldots \times \Pr(h_m)}}$$
$$= \mathfrak{pd}_K(h_1, E) \times \ldots \times \mathfrak{pd}_K(h_n, E) \times \frac{\mathfrak{pd}_K(h_1, \ldots, h_n|E)}{\mathfrak{pd}_K(h_1, \ldots, h_n)}$$

Proof of Theorem 6

$$\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, e_1 \cap \ldots \cap e_m) = \frac{\Pr(h_1 \cap \ldots \cap h_n \cap e_1 \cap \ldots \cap e_m)}{\Pr(h_1 \cap \ldots \cap h_n) \times \Pr(e_1 \cap \ldots \cap e_m)}$$
$$= \frac{\frac{\Pr(h_1 \cap \ldots \cap h_n \cap e_1 \cap \ldots \cap e_m)}{\Pr(h_1) \times \ldots \times \Pr(h_n) \times \Pr(e_1) \times \ldots \times \Pr(e_m)}}{\frac{\Pr(h_1 \cap \ldots \cap h_n)}{\Pr(h_1) \times \ldots \times \Pr(h_n)} \times \frac{\Pr(e_1 \cap \ldots \cap e_m)}{\Pr(e_1) \times \ldots \times \Pr(e_m)}}$$
$$= \frac{\mathfrak{pd}_K(h_1, \ldots, h_n, e_1, \ldots, e_m)}{\mathfrak{pd}_K(h_1, \ldots, h_n) \times \mathfrak{pd}_K(e_1, \ldots, e_m)}$$

Proof of Theorem 7 Let W be a set of possibilities and let $\mathcal{A}$ be some algebra over W. The elements of $\mathcal{A}$ are interpreted as propositions expressible in some suitable language $\mathcal{L}$ as specified in more detail in the appendix in connection with Theorem 2. The possibilities in W can be interpreted as models for $\mathcal{L}$. Let $e_0, \ldots, e_n, \ldots$ be a sequence of propositions of $\mathcal{A}$ that separates W, and let $e_i^w = e_i$ if $w \vDash e_i$ and $\overline{e_i}$ otherwise. Let Pr be a regular (or strict) probability function on $\mathcal{A}$. Let $\Pr^*$ be the unique probability function on the smallest σ-field $\mathcal{A}^*$ containing the field $\mathcal{A}$ satisfying $\Pr^*(A) = \Pr(A)$ for all $A \in \mathcal{A}$.
According to Theorem 2 there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all hypotheses $h_1, \ldots h_n$ and $h'_1, \ldots h'_m$ of $\mathcal{A}$.

1. If $w \vDash h_1 \cap \ldots \cap h_n$ and $w \vDash (\overline{h'_1 \cap \ldots \cap h'_m})$, then:
   $\exists k \forall l \geq k : [\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E_l^w) > \mathfrak{pd}_K(h'_1 \cap \ldots \cap h'_m, E_l^w)]$
2. If $w \vDash h_1 \cap \ldots \cap h_n \cap h'_1 \cap \ldots \cap h'_m$ and $h_1 \cap \ldots \cap h_n \vDash h'_1 \cap \ldots \cap h'_m$ but $h'_1 \cap \ldots \cap h'_m \nvDash h_1 \cap \ldots \cap h_n$, then:
   $\exists k \forall l \geq k : [\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E_l^w) > \mathfrak{pd}_K(h'_1 \cap \ldots \cap h'_m, E_l^w)]$
   where $E_l^w = \bigcap_{0 \leq i \leq l} e_i^w$.

Corollary 8 implies that:

$$\begin{aligned} &\mathfrak{pd}_K(h_1 \cap \ldots \cap h_n, E_l^w) > \mathfrak{pd}_K(h'_1 \cap \ldots \cap h'_m, E_l^w) \\ \Leftrightarrow &\frac{\mathfrak{pd}_K(h_1, \ldots, h_n, E)}{\mathfrak{pd}_K(h_1, \ldots, h_n)} > \frac{\mathfrak{pd}_K(h'_1, \ldots, h'_m, E)}{\mathfrak{pd}_K(h'_1, \ldots, h'_m)} \end{aligned}$$

Hence, there is a $W' \subseteq W$ with $\Pr^*(W') = 1$ so that the following holds for every $w \in W'$ and all hypotheses $h_1, \ldots h_n$ and $h'_1, \ldots h'_m$ of $\mathcal{A}$.

1. If $w \vDash h_1 \cap \ldots \cap h_n$ and $w \vDash (\overline{h'_1 \cap \ldots \cap h'_m})$, then:
   $\exists k \forall l \geq k : \left[\frac{\mathfrak{pd}_K(h_1, \ldots, h_n, E_l^w)}{\mathfrak{pd}_K(h_1, \ldots, h_n)} > \frac{\mathfrak{pd}_K(h'_1, \ldots, h'_m, E_l^w)}{\mathfrak{pd}_K(h'_1, \ldots, h'_n)}\right]$.
2. If $w \vDash h_1 \cap \ldots \cap h_n \cap h'_1 \cap \ldots \cap h'_m$ and $h_1 \cap \ldots \cap h_n \vDash h'_1 \cap \ldots \cap h'_m$ but $h'_1 \cap \ldots \cap h'_m \nvDash h_1 \cap \ldots \cap h_n$, then:
   $\exists k \forall l \geq k : \left[\frac{\mathfrak{pd}_K(h_1, \ldots, h_n, E_l^w)}{\mathfrak{pd}_K(h_1, \ldots, h_n)} > \frac{\mathfrak{pd}_K(h'_1, \ldots, h'_m, E_l^w)}{\mathfrak{pd}_K(h'_1, \ldots, h'_n)}\right]$
   where $E_l^w = \bigcap_{0 \leq i \leq l} e_i^w$.

References

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

[2] Brössel, P. (2008) Theory Assessment and Coherence. *abstracta* 4, 57–71.

[3] Brössel, P. (2012). *Rethinking Bayesian Confirmation Theory—Steps Towards a New Theory of Confirmation.* Unpublished PhD-Thesis (Konstanz).

[4] Brössel, P. (2013a). The Problem of Measure Sensitivity Redux. *Philosophy of Science* 80 (3), 378–397.

[5] Brössel, P. (2013b). Assessing Theories: The Coherentist Approach. *Erkenntnis* 79, 593–623.

[6] Brössel, P. (2013c). Correlation and Truth. in: V. Karakostas, D. Dieks (eds.). *Recent Progress in Philosophy of Science: Perspectives and Foundational Problems.* Dordrecht: Springer, 41-54.

[7] Carnap, R. (1962). *The Logical Foundations of Probability.* (2nd. edition). Chicago: University of Chicago Press.

[8] Crupi, V. (2013). Confirmation. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy.* URL = <http://plato.stanford.edu/archives/win2013/entries/confirmation/>.

[9] Crupi, V., Tentori, K. (2012). A Second Look at the Logic of Explanatory Power (With two Novel Representation Theorems). *Philosophy of Science* 79, 365–385.

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

[11] Fitelson, B. (1996). Wayne, Horwich, and Evidential Diversity. *Philosophy of Science* 63, 652–660.

[12] Fitelson, B. (1999). The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity. *Philosophy of Science* 66, S362–S378.

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

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

[15] Gaifman, H., Snir, M. (1982). Probabilities Over Rich Languages, Testing, and Randomness. *Journal of Symbolic Logic* 47, 495–548.

[16] 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.

[17] Good, I. J. (1983). *Good Thinking: The Foundations of Probability and Its Applications.* University of Minnesota Press.

[18] Hawthorne J (2014). Inductive logic. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), URL = <http://plato.stanford.edu/archives/sum2014/entries/logic-inductive/>.

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

[20] Horwich, P. (1982). *Probability and Evidence.* Cambridge: Cambridge University Press.

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

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

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

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

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

[26] Mackie, J. (1969). The Relevance Criterion of Confirmation. *The British Journal for the Philosophy of Science* 20, 27–40.

[27] McGrew, T. (2003). Confirmation, Heuristics, and Explanatory Reasoning. *The British Journal for the Philosophy of Science* 54, 553–567.

[28] Milne, P. (1996). $log[p(h/eb)/p(h/b)]$ is the one True Measure of Confirmation. *Philosophy of Science* 63, 21–26.

[29] Myrvold, W. (1996). Bayesianism and Diverse Evidence: A reply to Andrew Wayne. *Philosophy of Science* 63, 661–665.

[30] Myrvold, W. (2003). A Bayesian Account of the Virtue of Unification. *Philosophy of Science* 70, 399–423.

[31] Olsson, E. (2002). What Is the Problem of Coherence and Truth? *The Journal of Philosophy* 99, 246–272.

[32] Pollard, S. (1999). Milne's Measure of Confirmation. *Analysis* 59, 335–337.

[33] Popper, K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

[34] Schervish, M., Seidenfeld, T. (1990). An Approach to Consensus and Certainty with Increasing Evidence. *Journal of Statistical Planning and Inference* 25, 401–414.

[35] Schlesinger, G. (1995). Measuring Degrees of Confirmation. *Analysis* 55, 208–212.

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

[37] Shogenji, T. (2007). Why Does Coherence Appear Truth-Conducive? *Synthese* 157, 361–372.

[38] Schupbach, J. (2011a). Comparing Probabilistic Measures of Explanatory Power. *Philosophy of Science* 78, 813–829.

[39] Schupbach, J. (2011b). New Hope for Shogenji's Coherence Measure. *The British Journal for the Philosophy of Science* 62, 125–142.

[40] Schupbach, J. (2005). On a Bayesian Analysis of Unification. *Philosophy of Science* 72, 594–607.

[41] Schupbach, J., Sprenger, J. (2011). The Logic of Explanatory Power. *Philosophy of Science* 78, 105–127.

[42] Wayne, A. (1995). Bayesianism and Diverse Evidence. *Philosophy of Science* 62, 111–121.

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