

Reasoning with False Evidence

C. Günth¹

¹Institute of Philosophy,
Karlsruher Institute of Technology

Conference on Formal Models of Scientific Inquiry, 2017

Outline

- 1 Motivation
- 2 Veritistic Epistemology
- 3 Confirmation As Justification
- 4 Excursion - Higher Order Evidence
- 5 Simulations
- 6 Results

Outline

- 1 Motivation
- 2 Veritistic Epistemology
- 3 Confirmation As Justification
- 4 Excursion - Higher Order Evidence
- 5 Simulations
- 6 Results

Evidence - Fallibility and Role in Scientific Reasoning

- Empirical evidence is fallible and many of our present or past evidential beliefs have actually been false.
- Our ability to correctly infer the truth or falsity of a hypothesis depends on whether our body of evidence is correct.

Main Question

How can we reason with false evidence and reliably infer hypotheses?

Outline

- 1 Motivation
- 2 Veritistic Epistemology
- 3 Confirmation As Justification
- 4 Excursion - Higher Order Evidence
- 5 Simulations
- 6 Results

Veritistic Valuable Inferences

Formal framework: Veritistic Social Epistemology according to Goldman [2003]

- Beliefs such as knowledge, error and ignorance possess fundamental veritistic value.
 - $V\text{-Value}(\text{Knowledge})=1$
 - $V\text{-Value}(\text{Ignorance})=0.5$
 - $V\text{-Value}(\text{Error})=0$
- Practices possess a veritistic value insofar as they promote the acquisition of fundamental veritistic value.

Main Question Rephrased

Given a body of evidence which is partly incorrect, is there a veritistic valuable practice of inferring hypotheses?

Veritistic Indicators

Definition

A measure I is a veritistic indicator, if I is used in a doxastic practice P , in order to indicate how to change a belief system.

- The reliability of a veritistic indicator I is assessed via the veritistic value of the associated practice P .
- Here, P is a statistic test with $H_0 \equiv \neg h$ and test statistic $I(h)$.
 $(1 - \alpha) = P(I(h) \notin \omega | \neg h)$ and $(1 - \beta) = P(I(h) \in \omega | h)$
measure the veritistic value of this practice.

Main Question Rephrased

Given a body of evidence which is partly incorrect, is there a reliable veritistic indicator for the truth of a hypothesis?

Veritistic Debate Analyses

Results and limitations of Betz [2015]:

- Evidence accumulation as well as argumentation improves the reliability of justification as a veritistic indicator.
- A totally correct body of evidence is assumed. Yet, in real debates, the body of evidence often contains false evidence claims.
- Only justification is considered as a veritistic indicator. Yet, in real debates, especially in scientific ones, beliefs are often formed according to some notion of relevance confirmation.

Main Question Rephrased

Given a body of evidence which is partly incorrect, is confirmation a reliable veritistic indicator?

Outline

- 1 Motivation
- 2 Veritistic Epistemology
- 3 Confirmation As Justification**
- 4 Excursion - Higher Order Evidence
- 5 Simulations
- 6 Results

Degrees Of Justification

Formal model of complex argumentation: Theory of dialectical structures, see Betz [2010]

- A dialectical structure, τ , consists of theses and deductive valid arguments possibly attacking and supporting one another.
- A truth-value assignment to propositions in τ is called a position. If truth values are assigned to all propositions in τ , then the position is called complete, otherwise partial.
- A complete position p is dialectically consistent if and only if
(i) it assigns complementary truth values to a sentence and its negation
(ii) it considers conclusions of arguments with true premisses as true.

Degrees Of Justification

Let p and q be partial positions within a dialectical structure τ . Let σ be the number of complete and dialectical consistent positions in τ and σ_p the number of complete and dialectical positions extending p .

Definitions

Degree of justification of p : $DOJ_\tau(p) = \frac{\sigma_p}{\sigma}$

Conditional degree of justification of p given q : $DOJ_\tau(p|q) = \frac{\sigma_{p\&q}}{\sigma_q}$

Confirmation As Justification

According to Betz [2012], it holds that *DOJ* satisfy the probability axioms of Kolmogorov. Due to this fact, Bayesian confirmation theories can be spelled out using *DOJ*:

$$C_B(h, e) \rightarrow C_{DOJ}(h, e)$$

To choose some confirmation measure out of the multitude of discussed confirmation measures, the following condition is imposed:

Condition on Confirmation Measures

There has to be good reasons to consider $C_{DOJ}(h, e)$ as an extension of the concept of deductive entailment. Good reasons are those given in Crupi et al. [2007], Crupi and Tentori [2013].

Confirmation Measures

Confirmation measures, which satisfy this condition, are the following ones. (See the appendix for explicit definitions.)

- $DOJ(h|e)$
- $Z_{DOJ}(h, e)$, defined in Crupi and Tentori [2010].
- $F_{DOJ}(h, e)$, defined in Kemeny and Oppenheim [1952] and Fitelson [2004]

$DOJ(h|e)$ is an absolute confirmation measure, $Z_{DOJ}(h, e)$ and $F_{DOJ}(h, e)$ are relevance confirmation measures.

In order to account for non-contingent hypotheses, $Z_{DOJ}(h, e)$ is extended:

$$Z'_{DOJ}(h, e) = \begin{cases} Z_{DOJ}(h, e), & 0 < DOJ(h) < 1 \\ 1, & DOJ(h) = 1 \\ -1, & DOJ(h) = 0 \end{cases}$$

Outline

- 1 Motivation
- 2 Veritistic Epistemology
- 3 Confirmation As Justification
- 4 Excursion - Higher Order Evidence**
- 5 Simulations
- 6 Results

Preliminary Notion of HOE

Definition

A body of evidence is considered as first-order evidence and a statement about first-order evidence is considered as HOE.

Examples

In our framework, examples for HOE claims are statements about the amount and correctness of first-order evidence and properties of the argumentative structure into which first-order evidence is embedded.

Another Main Question

Does HOE allow us to estimate the reliability of confirmation as veritistic indicator for the truth of a hypothesis?

Another Notion of HOE

If (i) the answer to the above question is yes and (ii) inferring the truth of a hypothesis from its confirmation is understood as a fundamental inductive mode of reasoning, then a similar understanding of HOE can be found in Christensen [2010]:

- “[Higher order evidence] is information that effects what beliefs an agent (even an ideal agent) is epistemically rational in forming.”
- “[Higher order evidence] seems, at first blush, to be evidence of a peculiar sort. For one thing, its evidential import is often agent relative. For another, respecting it can apparently force an agent to fall short in certain ways, by having beliefs that fail to respect logic or basic inductive support relations.”

Outline

- 1 Motivation
- 2 Veritistic Epistemology
- 3 Confirmation As Justification
- 4 Excursion - Higher Order Evidence
- 5 Simulations**
- 6 Results

The Whole Simulation Run

- To answer the above questions, a veritistic analysis is performed over 1000 debates drawn from Betz [2013].
- Each debate contains 20 propositions and consists of a series of dialectical structures. Two adjacent dialectical structures differ only in one argument.

Example

Dialectical structure τ_1 : $(12 \wedge 19 \implies 10) \wedge (\neg 18 \wedge \neg 10 \implies 20) \wedge$
 $(\neg 13 \wedge \neg 8 \implies 6) \wedge (\neg 10 \wedge \neg 1 \implies \neg 11) \wedge (\neg 19 \wedge 15 \implies \neg 16) \wedge$
 $(5 \wedge 6 \implies 17) \wedge (\neg 5 \wedge 13 \implies 1) \wedge (\neg 15 \wedge \neg 6 \implies 7) \wedge$
 $(\neg 12 \wedge 15 \implies 2) \wedge (\neg 17 \wedge 18 \implies 7)$

Two adjacent dialectical structures: τ_1 and τ_2 with $\tau_2 = \tau_1 \wedge$
 $(\neg 8 \wedge \neg 7 \implies \neg 5)$

Simulations For a Single Debate

Choosing the truth

According to Betz [2015], for each debate, the truth is randomly chosen amongst all complete positions which are finally dialectical consistent.

Example

Complete position which is finally consistent, i.e. a candidate for the truth: $\neg 1 \wedge \neg 2 \wedge \neg 3 \wedge \neg 4 \wedge \neg 5 \wedge 6 \wedge 7 \wedge 8 \wedge 9 \wedge 10 \wedge \neg 11 \wedge \neg 12 \wedge \neg 13 \wedge 14 \wedge \neg 15 \wedge 16 \wedge 17 \wedge \neg 18 \wedge \neg 19 \wedge \neg 20$

Simulations For a Single Debate

Evidence accumulation

For each dialectical structure and each hypothesis, a maximal body of evidence is generated:

- 1 A pseudo-random sample of the truth is generated and the hypothesis under consideration is removed.
- 2 For a certain ratio of sentences, truth-value assignments are reversed.
- 3 A pseudo-random sample of this partial position is generated.

As evidence accumulates, more and more evidence claims of this maximal body of evidence are taken under consideration.

Simulations For a Single Debate

Evidence accumulation

Example

Random sample of the truth with sentence 9 removed:

$\neg 13 \wedge \neg 1 \wedge 17 \wedge \neg 19 \wedge \neg 4 \wedge \neg 8 \wedge 14 \wedge 5 \wedge 10 \wedge$

$\neg 15 \wedge \neg 2 \wedge \neg 11 \wedge \neg 18 \wedge \neg 12 \wedge \neg 3 \wedge 7 \wedge 6, \wedge \neg 20 \wedge 16$

Reversion of the fourth endmost truth-value assignments:

$\neg 7 \wedge \neg 6 \wedge 20 \wedge \neg 16$

Random sample of this modified partial position:

$\neg 13 \wedge \neg 19 \wedge \neg 16 \wedge 5 \wedge 14 \wedge 20 \wedge \neg 7 \wedge \neg 4 \wedge 17 \wedge \neg 2 \wedge \neg 6 \wedge \neg 15 \wedge$

$10 \wedge \neg 11 \wedge \neg 12 \wedge \neg 8 \wedge \neg 1 \wedge \neg 18 \wedge \neg 3$

Accumulation of evidence: $\{\{\}, \{\neg 13\}, \{\neg 13 \wedge \neg 19\}, \{\neg 13 \wedge \neg 19 \wedge \neg 16\}, \dots, \{\neg 13 \wedge \neg 19 \wedge \neg 16 \wedge 5 \wedge 14 \wedge 20 \wedge \neg 7 \wedge \neg 4 \wedge 17 \wedge \neg 2 \wedge \neg 6 \wedge \neg 15 \wedge 10 \wedge \neg 11 \wedge \neg 12 \wedge \neg 8 \wedge \neg 1 \wedge \neg 18 \wedge \neg 3\}\}$

Simulations For a Single Debate

Calculating degrees of confirmation

For each hypothesis of a certain debate, the degree of confirmation is calculated given

- a certain amount and correctness of evidence claims
- a certain inferential density¹

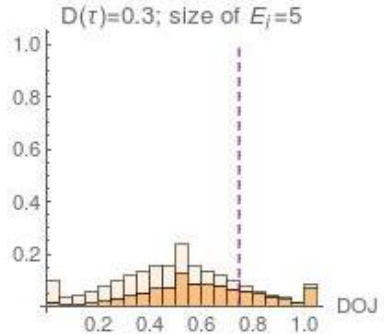
In the following, “size of E_i ” or $|e|$ refers to the amount of evidence claims, $1 - acc$ to the ratio of false evidence claims and $D(\tau)$ to the inferential density.

¹Roughly, the more arguments a dialectical structure hosts, the higher its inferential density.

The Whole Simulation Run

Confirmation histograms

The analysis is performed over an ensemble of 1000 debates. The result can be visualized as a histogram regarding degrees of confirmation. (The bars in darker (lighter) shading represent the fraction of true (false) hypotheses within a certain interval of confirmation.)



The Whole Simulation Run

Reliability of confirmation as veritistic indicator

According to James [2006], a statistic test can be performed with the help of each confirmation histogram:

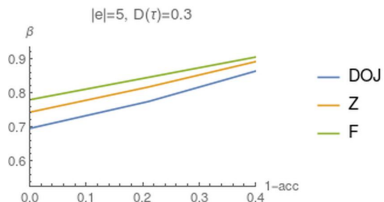
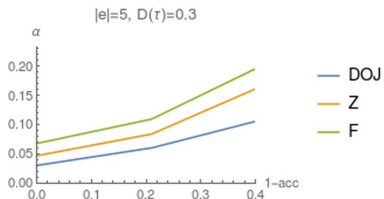
- The hypotheses to be tested is $\neg h$ and it is tested against the alternative hypotheses h
- The test statistic is a confirmation measure, namely $DOJ(h|e)$, $Z'(h, e)$ or $F(h, e)$
- The critical region ω is chosen in such a way that (i) $1 - \beta = 0.25$ or (ii) $\alpha = 0.05$ with $1 - \beta = P(C(h, e) \in \omega|h)$ and $\alpha = P(C(h, e) \in \omega|\neg h)$
- (i) α or (ii) β is calculated

Outline

- 1 Motivation
- 2 Veritistic Epistemology
- 3 Confirmation As Justification
- 4 Excursion - Higher Order Evidence
- 5 Simulations
- 6 Results**

Reliability of Confirmation as Veritistic Indicator

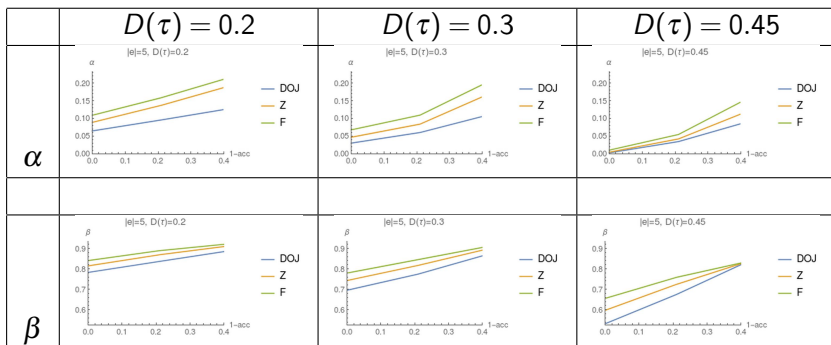
Influence of the ratio of false evidence claims



- 1 Absolute confirmation is generally a more reliable veritistic indicator than relevance confirmation.
- 2 As the ratio of false evidence claims increases, the reliability of confirmation as veritistic indicator decreases. This is true for absolute as well as relevance confirmation.

Reliability of Confirmation as Veritistic Indicator

Influence of inferential density - graphically



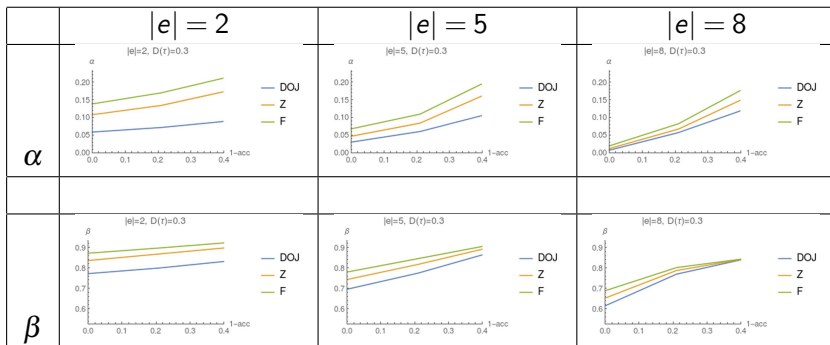
Reliability of Confirmation as Veritistic Indicator

Influence of inferential density - in words

- 1 Absolute confirmation is generally a more reliable veritistic indicator than relevance confirmation.
- 2 As the inferential density increases, the reliability of confirmation as veritistic indicator increases. This is true for all ratios of false evidence claims.
- 3 The distance between reliability - as measured by $1 - \alpha$ - of absolute confirmation and relevance confirmation decreases for increasing inferential density.
- 4 The distance between reliability - as measured by $1 - \beta$ - of absolute confirmation and relevance confirmation increases for increasing inferential density, as long as inferential density and ratio of false evidence claims are not too high.

Reliability of Confirmation as Veritistic Indicator

Influence of accumulation of evidence - graphically



Reliability of Confirmation as Veritistic Indicator

Influence of accumulation of evidence - in words

- 1 Absolute confirmation is generally a more reliable veritistic indicator than relevance confirmation.
- 2 As more and more evidence is accumulated, the reliability of confirmation as veritistic indicator increases. This is not true for absolute confirmation and 40% false evidence claims.
- 3 The distance between reliability of absolute confirmation and relevance confirmation decreases as more and more evidence is accumulated.

Summary

- 1 Given a body of evidence which is partly incorrect, absolute confirmation is a more reliable veritistic indicator for the truth of a hypothesis than relevance confirmation.
- 2 Higher order evidence allows us to estimate the reliability of confirmation as veritistic indicator for the truth of a hypothesis.
 - 1 As the ratio of false evidence claims increases, the reliability decreases.
 - 2 As the inferential density increases, the reliability increases.
 - 3 As more and more evidence is accumulated, the reliability increases. This is not true for absolute confirmation and 40% false evidence claims.

Outlook

Given a body of evidence which is partly incorrect, the reliability of confirmation as veritistic indicator for the truth of a hypothesis has to be improved. Therefore, the following actions are going to be performed:

- 1 Revision of the body of evidence according to some revision rule, as for example “Remove claims from the body of evidence in such a way that its coherence is maximized.”
- 2 Modification of the confirmation measures according to Crupi et al. [2008].

Veritistic Practice Under Consideration - Statistic Tests

Here, a statistic test is considered with null hypotheses $H_0 \equiv \neg h$ and test statistic $I(h)$. It is a practice governed by the following inferential rule:

Statistic Test Rule

If $I(h)$ falls into the critical region of $H_0 \equiv \neg h$, then accept h .

There are two kinds of error probabilities of such inferences:

Error probabilities of this Statistic Test

$\alpha = P(I(h) \in \omega | \neg h)$ and $\beta = P(I(h) \notin \omega | h)$ with ω as critical region of $H_0 \equiv \neg h$

$(1 - \alpha)$ and $(1 - \beta)$ measure the veritistic value of this practice, respectively the reliability of the test statistic as veritistic indicator.

Qualitative Notions of Confirmation

Let h be a hypothesis, e some evidence and B a subjective probability function.

Definitions

Absolute Confirmation.

e confirms $h \iff B(h|e) > B(\neg h|e)$

e dis-confirms $h \iff B(h|e) < B(\neg h|e)$

Relevance Confirmation.

e confirms $h \iff B(h|e) > B(h)$

e dis-confirms $h \iff B(h|e) < B(h)$

In order to obtain a quantitative notion of confirmation, a confirmation measure, $C_B(h, e)$, has to be specified.

Quantitative Notion of Confirmation Spelled Out

$$\bullet F_B(h, e) \equiv \begin{cases} L_B(h, e), & 0 < B(h|e) < 1 \\ 1, & B(h|e) = 1 \wedge B(e) \neq 0 \text{ with} \\ -1, & B(h|e) = 0 \end{cases}$$

$$L_B(h, e) = \frac{B(e|h) - B(e|\neg h)}{B(e|h) + B(e|\neg h)}$$

$$\bullet Z'_B(h, e) = \begin{cases} Z_B(h, e), & 0 < B(h) < 1 \\ 1, & B(h) = 1 \quad \text{with} \\ -1, & B(h) = 0 \end{cases}$$

$$Z_{DOJ}(h, e) \equiv \begin{cases} \frac{B(h|e) - B(h)}{1 - B(h)}, & B(h|e) \geq B(h) \\ \frac{B(h|e) - B(h)}{B(h)}, & B(h|e) < B(h) \end{cases}$$

References I

Gregor Betz. *Theorie dialektischer Strukturen*. Klostermann, Frankfurt am Main, 2010. ISBN 978-3-465-03629-6.

Gregor Betz. On Degrees of Justification. *Erkenntnis*, 77(2): 237–272, September 2012. ISSN 0165-0106, 1572-8420. doi: 10.1007/s10670-011-9314-y. URL <http://link.springer.com/10.1007/s10670-011-9314-y>.

Gregor Betz. *Debate dynamics how controversy improves our beliefs*. Springer, Dordrecht; New York, 2013. ISBN 978-94-007-4599-5 978-94-007-4598-8.

References II

Gregor Betz. Truth in evidence and truth in arguments without logical omniscience. *The British Journal for the Philosophy of Science*, page axv015, 2015. ISSN 0007-0882, 1464-3537. doi: 10.1093/bjps/axv015. URL <http://bjps.oxfordjournals.org/content/early/2015/03/31/bjps.axv015>.

David Christensen. Higher-order evidence. *Philosophy and Phenomenological Research*, 81(1):185–215, 2010.

Vincenzo Crupi and Katya Tentori. Irrelevant Conjunction: Statement and Solution of a New Paradox. *Philosophy of Science*, 77(1):1–13, January 2010. ISSN 0031-8248. doi: 10.1086/650205. URL <http://www.journals.uchicago.edu/doi/citedby/10.1086/650205>.

References III

- Vincenzo Crupi and Katya Tentori. Confirmation as partial entailment: A representation theorem in inductive logic. *Journal of Applied Logic*, 11(4):364 – 372, 2013. ISSN 1570-8683. doi: <http://dx.doi.org/10.1016/j.jal.2013.03.002>. URL <http://www.sciencedirect.com/science/article/pii/S1570868313000128>. Combining Probability and Logic: Papers from Prolog 2011.
- Vincenzo Crupi, Katya Tentori, and Michel Gonzalez. On bayesian measures of evidential support: Theoretical and empirical issues*. *Philosophy of Science*, 74(2):229–252, 2007. ISSN 00318248, 1539767X. URL <http://www.jstor.org/stable/10.1086/520779>.

References IV

- Vincenzo Crupi, Roberto Festa, and Tommaso Mastropasqua.
Bayesian confirmation by uncertain evidence: A reply to huber [2005]. *ResearchGate*, 59(2), jun 2008. ISSN 0007-0882. doi: 10.1093/bjps/axn008. URL https://www.researchgate.net/publication/31171207_Bayesian_Confirmation_by_Uncertain_Evidence_A_Reply_to_Huber_2005.
- Branden Fitelson. Two technical corrections to my coherence measure. 2004. URL <http://fitelson.org/coherence2.pdf>.
- Alvin I Goldman. *Knowledge in a social world*. Clarendon Press ; Oxford University Press, Oxford; New York, 2003. ISBN 978-0-19-823777-8 978-0-19-823820-1.

References V

F James. *Statistical methods in experimental physics*. World Scientific, Hackensack, NJ, 2006. ISBN 978-981-256-795-6 978-981-270-527-3. OCLC: 85207490.

John G. Kemeny and Paul Oppenheim. Degree of factual support. *Philosophy of Science*, 19(4):307–324, 1952.