



CHICAGO JOURNALS



---

The Independence Condition in the Variety-of-Evidence Thesis

Author(s): François Claveau

Reviewed work(s):

Source: *Philosophy of Science*, Vol. 80, No. 1 (January 2013), pp. 94-118

Published by: [The University of Chicago Press](#) on behalf of the [Philosophy of Science Association](#)

Stable URL: <http://www.jstor.org/stable/10.1086/668877>

Accessed: 30/01/2013 09:13

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press and Philosophy of Science Association are collaborating with JSTOR to digitize, preserve and extend access to *Philosophy of Science*.

<http://www.jstor.org>

# The Independence Condition in the Variety-of-Evidence Thesis

François Claveau\*†

---

The variety-of-evidence thesis has been criticized by Bovens and Hartmann. This article points to two limitations of their Bayesian model: the conceptualization of unreliable evidential sources as randomizing and the restriction to comparing full independence to full dependence. It is shown that the variety-of-evidence thesis is rehabilitated when unreliable sources are reconceptualized as systematically biased. However, it turns out that allowing for degrees of independence leads to a qualification of the variety-of-evidence thesis: as Bovens and Hartmann claimed, more independence does not always imply stronger confirmation.

---

**1. Introduction.** Seeking a variety of evidence for a hypothesis is standard practice in science, as well as in normal life. The members of the OPERA Collaboration, for instance, appealed to the value of evidential variety when they disclosed their measurement of neutrinos apparently traveling faster than light: “While OPERA researchers will continue their studies, we are also looking forward to independent measurements to fully assess the nature of this observation” (Istituto Nazionale di Fisica Nucleare 2011).

Evidential variety is also prized in economics. For example, it is commonplace in labor economics for a causal hypothesis to be seen as more strongly supported if it can rely not only on macrodata evidence but also on microdata evidence. If the hypothesis under consideration is ‘the relatively long duration of unemployment benefits in France is a cause of its relatively high unemployment rate’, the proposition ‘there is a positive statistical association between average duration of benefits and unemployment rates

Received April 2012; revised July 2012.

\*To contact the author, please write to: Faculty of Philosophy, Erasmus University, Rotterdam; e-mail: [claveau@fwb.eur.nl](mailto:claveau@fwb.eur.nl).

†I thank Stephan Hartmann, Kevin Hoover, Conor Mayo-Wilson, Luca Moretti, and Julian Reiss for their helpful comments, James Kelleher for English proofreading, and the SSHRC (767-2009-0001) for financial support.

Philosophy of Science, 80 (January 2013) pp. 94–118. 0031-8248/2013/8001-0008\$10.00  
Copyright 2013 by the Philosophy of Science Association. All rights reserved.

among industrial countries' (macrodata evidence) will be interpreted as supporting the hypothesis, but the support will be even higher if the evidential elements also include the proposition 'the average length of an unemployment spell increased in Austria for the category of job seekers affected by the 1989 reform of benefits duration' (microdata evidence).<sup>1</sup>

The widespread quest for evidential variety can be justified by what Bayesians call the variety-of-evidence thesis.

**Variety-of-evidence thesis.** *Ceteris paribus*, the strength of confirmation of a hypothesis by an evidential set increases with the diversity of the evidential elements in that set.

Some Bayesians maintain that this thesis could be given a formal proof once its key terms—the *ceteris paribus* clause, confirmation, variety—are properly defined. In seeking this proof, the most popular interpretation of variety has been to equate it to a measure of independence among evidential elements.<sup>2</sup> The intuitive idea behind the proposals of Earman (1992) and Howson and Urbach (1993) is that an evidential set is varied to the extent that each element  $e_i$  is not made significantly more likely by learning other elements in the set—the extreme case being full probabilistic independence between  $e_i$  and any conjunct of the other elements.

It turns out that one runs into problems in trying to prove the variety-of-evidence thesis using such a measure of independence. It is indeed clear from a measure introduced by Myrvold (1996) and recently labeled “focused correlation” by Wheeler (2009) that, in order to prove the variety-of-evidence thesis using the most popular interpretation of variety, one must either assume that the hypothesis entails the evidence, which would be forgetting the role of auxiliary hypotheses, or smuggle into the *ceteris paribus* clause the measure of independence conditional on the hypothesis, which seems unwarranted.

In parallel to these developments, Bovens and Hartmann introduced another characterization of variety as *reliability* independence. Using this notion of independence, they challenged the belief that Bayesianism can prove the variety-of-evidence thesis. According to their model, “less varied evidence may indeed provide more confirmation to the hypothesis” (Bovens and Hartmann 2002, 47; 2003, 106).

1. These evidential propositions summarize, respectively, the results from OECD (2006, table 3.3) and Lalive, van Ours, and Zweimüller (2006).

2. In contrast, Horwich (1982, 1998) connected variety with the capacity of an evidential set to disconfirm alternative hypotheses. For discussions and criticisms, see Wayne (1995), Fitelson (1996), and Bovens and Hartmann (2003, 107).

Bovens and Hartmann use what seems to be a plausible understanding of variety: evidential elements for a given hypothesis are varied to the extent that they do not share potential reasons for being unreliable. For example, the ICARUS Collaboration (2012) was in a position to produce evidence for or against the hypothesis of faster-than-light neutrinos, evidence that was partially independent of the OPERA experiment. The independence is partial here because, while on the one hand, the two groups shared the same neutrino beams from CERN (making them share some potential reasons to be unreliable), on the other hand, they used different detectors—opening up the possibility that the ICARUS measurement is unbiased while the OPERA measurement is systematically biased due to a defect in the OPERA detector.<sup>3</sup>

This article takes a second look at Bovens and Hartmann's result. My primary concern is to assess whether their result should affect the status of the variety-of-evidence thesis as a guide to scientific practice. Endorsing their (plausible) interpretation of variety as reliability independence, I argue that two aspects of their model shed doubt on the relevance of their result for actual science. First, the *unreliable* sources in their model are not like unreliable sources in actual science (i.e., their unreliable sources are randomly biased while systematic bias is far more likely to be the issue). I show, in section 4, that the variety-of-evidence thesis is rescued when the model is slightly modified to capture unreliability as systematic bias. Second, their model, and my first modification to it, contrasts full independence to full dependence, while variety in the variety-of-evidence thesis is more a question of degrees of independence. In section 5, I extend the model to consider degrees of independence. I then show that the variety-of-evidence thesis, as Bovens and Hartmann initially claimed, is false.

**2. Bovens and Hartmann's Result.** In this section, I present a simplified version of Bovens and Hartmann's model and reproduce one of their results against the variety-of-evidence thesis.<sup>4</sup> The model uses three types of propositional variables:

- The hypothesis variable  $H = \{h, \neg h\}$ , where  $h$  stands for the proposition that the hypothesis of interest is true (e.g., 'some neutrinos can travel faster than light') and  $\neg h$  stands for its negation.

3. This is now the official explanation of the OPERA anomaly (CERN 2012).

4. Their model is slightly more complex because it adds another propositional variable to the three that I consider—i.e., the 'testable consequence'  $C$  (Bovens and Hartmann 2003, 89–90). Since I focus on the issue of independent reliability—not on the issue of independent testable consequences—this addition is superfluous.

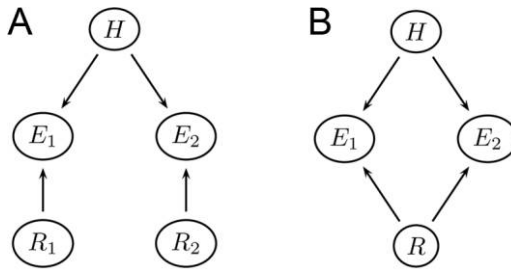


Figure 1. Two cases of partially reliable evidential sources. *A*, Independent reliability; *B*, shared reliability.

- The evidential variable  $E_i = \{e_i, \neg e_i\}$ , where  $e_i$  stands for a positive report regarding  $h$ , that is, a report to the effect that a testable consequence of  $h$  holds (e.g., ‘the measured velocity of the neutrinos in this experiment is higher than the speed of light’) and  $\neg e_i$  stands for a negative report.
- The reliability variable  $R_i = \{r_i, \neg r_i\}$ , where  $r_i$  stands for the proposition that the evidential source  $i$  (the one having as output  $E_i$ ) is reliable and  $\neg r_i$  stands for the proposition that the source is unreliable.

Two joint probability distributions are constructed over the set of variables  $\{H, E_1, E_2, R_1, R_2\}$ . The assumed probabilistic independencies among the variables can be read off the Bayesian networks in figure 1 by using the  $d$ -separation criterion (Pearl 1988, 117–18).<sup>5</sup>

The probability distribution  $P_I(\cdot)$  associated with the network in figure 1A is meant to capture the idea that two sources are *reliability independent* (i.e.,  $R_1 \perp\!\!\!\perp R_2$ ). The probability distribution  $P_S(\cdot)$  associated with figure 1B captures the other extreme when evidential sources have *fully shared reliabilities* (i.e.,  $R_1 = R_2 = R$ ). The notion of variety modeled here is thus:

**Reliability independence.** Two evidential elements are independent if their reliabilities are independent.

5. The two distributions share the condition  $H \perp\!\!\!\perp R_1, R_2, R$ —i.e., before learning the evidential report  $E_i$ , learning that the associated evidential source is reliable or not has no bearing on the strength of belief in the hypothesis (and vice versa). For fig. 1A, we also have  $E_i \perp\!\!\!\perp E_j, R_j | H$  for  $i \neq j$ , which means that once the realization of  $H$  is known, learning the realization of  $E_j$  or  $R_j$  for  $j \neq i$  is not relevant to the probability distribution of  $E_i$ . A similar condition for fig. 1B is  $E_i \perp\!\!\!\perp E_j | H, R$  for  $i \neq j$ , which means that, in this case, one needs to condition on the reliability variable, too, in order for the two evidential reports to be irrelevant to each other. These conditions do not universally apply to what can be considered evidential elements for a hypothesis (e.g., Wheeler and Scheines 2011, fig. 3). This fact must be kept in mind in interpreting the result of Bovens and Hartmann as well as my results.

The joint probability distributions are further specified. The root variables  $H$  and  $R_i$  are given prior probabilities:

$$P(h) = h_0 \quad \text{and} \quad P(r_i) = \rho_i, \quad (1)$$

where  $h_0$  and  $\rho_i$  are parameters strictly between 0 and 1. The prior degree of belief that the hypothesis is true is thus  $h_0$ , and  $\rho_i$  is the prior degree of belief that the evidential source  $i$  is reliable. For compactness, I will write  $\bar{h}_0$  for the prior probability that the hypothesis is false ( $1 - h_0$ ) and  $\bar{\rho}_i$  for the prior probability that source  $i$  is unreliable ( $1 - \rho_i$ ).

What remains to be spelled out is how the evidential variable  $E_i$  varies with its parents. It is assumed that, when a source is reliable, the evidential report is a perfect truth tracker:

$$P(e_i|h, r_i) = 1 \quad \text{and} \quad P(e_i|\neg h, r_i) = 0 \quad \text{for} \quad i = \{1, 2\}. \quad (2)$$

That is, when the hypothesis is true, a reliable evidential source will give a positive report; when the hypothesis is false, such a source will give a negative report.

What happens when the source is unreliable? To specify this case, Bovens and Hartmann rely on the following intuition:

**Irrelevance of an unreliable source.** If one knows for sure that a given source is unreliable ( $R_i = \neg r_i$ ), the report coming from this source ( $e_i$  or  $\neg e_i$ ) should not have any effect on the degree of belief in the hypothesis  $h$ .

This can be written

$$P(h|e_i, \neg r_i) = P(h|\neg e_i, \neg r_i) = P(h|\neg r_i). \quad (3)$$

In other words, an unreliable source gives garbage information regarding the truth of the hypothesis. Upon learning the information from an unreliable source, the agent makes no updating to the subjective probability of the hypothesis.

Note that this is not the only plausible interpretation of ‘unreliable source’. The interpretation clashes, in particular, with the idea that unreliability might be due to calibration issues. Taking again the OPERA experiment as an example, some early critics claimed that the anomalous result might be due to a problem with clock synchronization (Contaldi 2011). The estimated time of travel would be systematically below the actual time because the clock at the end of the tube clicked slightly later than the clock at the beginning.<sup>6</sup> With this type of unreliability, it is possible to undo the bias: given the estimated value and given the bias, one can retrieve the actual travel time. Knowing, for

6. The OPERA researchers have indeed identified similar biases by now.

example, that the experimental setup is biased ( $\neg r_i$ ) in such a way that the estimated time is systematically underestimated by a factor  $b$ , obtaining a positive report  $e_i$  that the time of travel is  $t$  should matter for one's belief in the hypothesis because an unbiased estimate of the travel time can be retrieved by computing  $t/b$ . For such calibration problems, we thus have that  $P(h|e_i, \neg r_i) \neq P(h|\neg r_i)$ , which contradicts condition (3). In this article, I stick to interpretations of unreliability compatible with condition (3) and keep the calibration interpretation for future work. I will later give examples of reasons for unreliability that are compatible with condition (3).

Condition (3) implies what follows (proof in Bovens and Hartmann 2003; app. C.1):

$$P(e_i|h, \neg r_i) = P(e_i|\neg h, \neg r_i) =: \alpha_i. \tag{4}$$

Another way to express this condition is  $E_i \perp\!\!\!\perp H | \neg r_i$ .

Parameter  $\alpha_i$  is the last parameter of the model; it is the probability that the evidential report is positive, given that the source is unreliable. The probability that the evidential report is negative, given that the source is unreliable, is simply  $1 - \alpha_i =: \bar{\alpha}_i$ . Table 1 sums up how the realizations of  $E_i$  depend on the values taken by  $H$  and  $R_i$ .

Bovens and Hartmann offer a specific interpretation of  $\alpha_i$  in terms of a randomizing evidential source. I will later offer an alternative interpretation of this parameter, but let me first reproduce their result relative to the variety-of-evidence thesis. The probability of interest is the posterior belief in the hypothesis given two positive reports:  $P(h|e_1, e_2) = P^*(h)$ . For the two joint probability distributions  $P_I(\cdot)$  and  $P_S(\cdot)$ —that is, the distributions associated with the reliability-independent version (fig. 1A) and the shared-reliability version (fig. 1B)—this posterior can be written using the likelihood-ratio form

$$P^*(h) = \frac{h_0}{h_0 + h_0 L}, \quad \text{where } L = \frac{P(e_1, e_2 | \neg h)}{P(e_1, e_2 | h)}. \tag{5}$$

Which posterior is higher,  $P_I^*(h)$  or  $P_S^*(h)$ ? To turn this comparison into an assessment of the variety-of-evidence thesis, we need a plausible interpretation of the ceteris paribus condition of this thesis. Bovens and Hartmann impose restrictions that seem sufficient to meet the condition. First,

TABLE 1. PROBABILITY OF A POSITIVE REPORT GIVEN THE VALUES OF  $H$  AND  $R_i$

$P(e_i H, R_i)$	$r_i$	$\neg r_i$
$h$	1	$\alpha_i$
$\neg h$	0	$\alpha_i$

we want to rule out comparing hypotheses starting with unequal degrees of confirmation. We thus impose  $P_I(h) = P_S(h)$ . Second, we want the evidential sets to potentially differ in confirmatory strengths for no other reason than their relative variety. A *sufficient* condition for this goal is to require that all positive reports  $e_i$  in the independent-reliability and the shared-reliability versions have the same confirmatory strength for  $h$ —that is,  $P_I(h|e_i) = P_S(h|e_i)$  for  $i, j = \{1, 2\}$ . This condition holds if the different  $\alpha_i$  and  $\rho_i$  are reduced to only a single  $\alpha$  and a single  $\rho$  across the two models (Bovens and Hartmann 2003, 104).

Given this interpretation of the *ceteris paribus* condition, the likelihood ratios associated with the two posteriors  $P_I^*(h)$  and  $P_S^*(h)$  are (proof in app. A)

$$L_I = \frac{(\alpha\bar{\rho})^2}{(\rho + \alpha\bar{\rho})^2}, \quad (6)$$

$$L_S = \frac{\alpha^2\bar{\rho}}{\rho + \alpha^2\bar{\rho}}. \quad (7)$$

Since we assume that the prior probability of the hypothesis is the same in the two models, the variety-of-evidence thesis implies that  $P_I^*(h) > P_S^*(h)$  for all admissible parameter values—that is, we should have a higher confidence in the hypothesis if our two positive reports come from reliability-independent sources as compared to reliability-shared sources. This is equivalent to  $L_S > L_I$ . It turns out, however, that the inequality is reversed for some combinations of values of  $\alpha$  and  $\rho$  (proof in app. A):

$$P_I^*(h) > P_S^*(h) \Leftrightarrow .5 > \alpha\bar{\rho}. \quad (8)$$

Figure 2 divides the parameter space in two regions: the bigger white region where independence is more confirmatory and the gray region where shared reliability is more confirmatory.

What happens? Why is it sometimes better for confirmation to have no reliability independence rather than full independence? To understand this, it is crucial to see what shared reliability entails in the second version of the model, namely, that  $E_1$  is a truth teller if and only if  $E_2$  is a truth teller. As truth tellers,  $E_1$  and  $E_2$  will always give concordant reports. But when none of the evidential variables is a truth teller (i.e., when  $\neg r$ ), then each evidential variable has a probability  $\alpha$  of producing a positive report. It is crucial to recognize that this probability is not affected by the value the other evidential variable is realizing. It implies that when, and only when, they are unreliable,  $E_1$  and  $E_2$  might realize discordant reports. A second concordant report in the shared-reliability model thus contributes to confirmation in the following way: “we feel more confident that the instrument is not a randomizer and this increase in confidence in the reliability of the



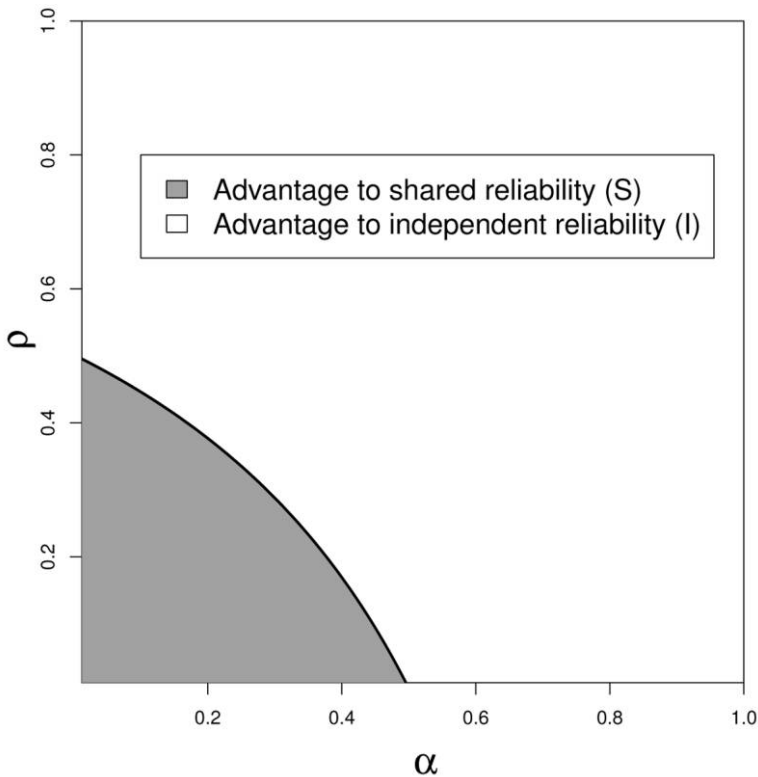


Figure 2. Parameter space showing when shared reliability is more confirmatory than independent reliability.

instrument benefits the confirmation of the hypothesis” (Bovens and Hartmann 2003, 98). The region of the parameter space where shared reliability is better is the one where this channel of ‘higher confirmation of  $h$  because higher confidence in the reliability of the source’ is the most effective. With low values of  $\alpha$ , it is unlikely that an unreliable source would output two positive reports; it is thus likely that the two positive reports come from a reliable source. With low values of  $\rho$ , the agent starts with little confidence in the source; there is both great room for improving confidence and little to be gained for the belief in the hypothesis from the direct effect of a positive report since such a report is not likely to be truth tracking.

**3. Questioning Bovens and Hartmann’s Result.** While the logic of this result is simple, its implications for the variety-of-evidence thesis are less clear. As Bovens and Hartmann (2003, 95n) recognize, the result of their model “does not apply to unreliable instruments that do not randomize.” It

thus seems that the champions of the variety-of-evidence thesis would have little to worry about if evidential sources were rarely as Bovens and Hartmann depict them to be. And it indeed seems to be the case that scientists do not think of their evidential sources in the manner depicted by the shared-reliability model.

For example, consider the macrodata evidence for the hypothesis that one cause of the relatively high French unemployment rate is its relatively long duration of unemployment benefits. To simplify, let us imagine that the macrodata evidence is Pearson's correlation coefficient between the legislated duration of unemployment benefits and the unemployment rate for a sample of industrial countries. There are many potential reasons why this coefficient would not be truth tracking (i.e., would be unreliable) with respect to the hypothesis of interest—for example, the correlation might not be a sign of causation from benefits to the unemployment rate because of the presence of a common cause, or perhaps the causal structure in France deviates substantially from the ones in the sampled countries.

Now imagine that I decide to compute the correlation coefficient twice. That is, I have the data for the two variables in my computer, and I use my favorite statistical software to compute the correlation twice (e.g., send the command `cor(Bd,U)` to R twice). It is reasonable to say that the two results would share a single reliability state: my second correlation coefficient would be reliable evidence for the hypothesis if and only if my first coefficient is also reliable evidence for the hypothesis. Does Bovens and Hartmann's shared-reliability situation come anywhere close to capturing how we think about these two results? Obviously not.

According to the model, the two results will always be in concordance if the procedure is reliable. This implication seems fine. But something strange must happen with my statistical software when the procedure is unreliable. In this case, the two results might be at odds. Furthermore, if I were to compute the coefficient again and again, I would necessarily get discordant results (provided  $\alpha$  is strictly between 0 and 1). To be sure, it is possible that my procedure is unreliable and is randomizing in this way. My correlation command might have been redefined such that it randomly outputs a number between  $-1$  and  $1$ . However, this is not what would normally be of concern. The reasons given above for why a correlation might be unreliable evidence for a specific causal claim will not bring about such randomness. If, for example, there is a confounding common cause, one would expect both coefficients to be identically affected.

Another way to see the problem with the model is to imagine that the data used to compute the correlation coefficients are known to be totally unrelated to French unemployment. Let us say that the two variables are the respective prices of two types of fish at the Grote Markt in Rotterdam. Since the correlation between these two variables is not tracking the truth of

the hypothesis about French unemployment, the source is unreliable for this hypothesis ( $R = \neg r$ ). Now the model tells us that, since the source is unreliable, the two computed correlations must be probabilistically independent. The agent thus starts with some strength of belief  $\alpha$  that the correlation between the fish prices is positive. She sends the correlation command and reads a first positive coefficient. Strangely, learning this first coefficient will not make her revise her belief about the probable value of the second coefficient; just before pressing Enter again on her computer, she will still believe to strength  $\alpha$  that the computer output will be positive.

The counterintuitiveness of Bovens and Hartmann's model is not an artifact of my specific choice of example. Take the complex experimental setup of the OPERA Collaboration. Imagine that the research team—before announcing that it had located biases in its experimental procedure—had rerun the experiment with the exact same setup (imagine this to be the case even though the exact same setup is a physical impossibility) and that the results had corroborated the initial measurement. What would have been the reaction of the scientific community? The model tells us that the new results should have been taken as evidence that the setup is truth tracking. But it seems more intuitive that these results would have been met with indifference. Scientists distrusted the first measurement because they believed that the experiment suffered from a systematic (yet unknown) bias. Concordant results from a second identical experiment could thus be explained away by saying that the systematic bias was again operating (as it should if experimenters were careful enough in reproducing the setup). To make some progress in the debate, OPERA researchers needed to find a way to decrease the strength of the belief in the existence of a systematic bias. Rerunning the exact same experiment over and over would not have achieved that.

The upshot of this discussion is that the result of Bovens and Hartmann, as it stands, should not worry scientists and philosophers very much, if at all. There is still room for an unqualified variety-of-evidence thesis when the sources of evidence resemble the ones in science, rather than the ones in the model.<sup>7</sup>

**4. First Modification to the Model.** Doubt has crept in: Can it not be shown that a more appropriate modeling of the evidential sources still results in a qualified variety-of-evidence thesis? In this section, I offer a negative answer to this question by making a single modification to the model of Bovens and Hartmann.

7. Hartmann (2008, 108) later wrote the following statement about his assumption regarding unreliability: "This way of modeling a partially reliable instrument is clearly a strong idealization, which will not hold in many cases."

Bovens and Hartmann's modeling choices are guided by a specific interpretation of the parameter  $\alpha$ : for them it means that an unreliable evidential source acts like a randomizer. This becomes clear in their discussion of witnesses as a special case of an evidential source: "So, we assume that if witnesses are not reliable, then they are like randomizers. It is as if they do not even look at the state of the world to determine whether the hypothesis is true, but rather flip a coin or cast a die to determine whether they will provide a report to the effect that the hypothesis is true" (Bovens and Hartmann 2003, 57). While they interpret the parameter as capturing a property of the evidential source, I would rather interpret it from the point of view of the agent: just knowing that the source  $i$  is unreliable,  $\alpha_i$  is the agent's degree of belief that the report of this source will be positive.

It turns out that Bovens and Hartmann could have modeled this epistemic fact in a different way. Such an alternative way specifies that an unreliable evidential source is systematically biased, not randomizing. I want to emphasize at the outset that systematically biased sources of the kind I will model do not cover all the potential kinds of unreliability in science. Obviously, they do not cover randomizing sources (if such sources exist). More important, they fail to encompass the miscalibrated sources previously mentioned in section 2.

I still think that what I model as 'systematically biased sources' capture important reasons why one can judge a source to be unreliable. Here are a few hints at these reasons without any claim to be comprehensive. One general class of cases comprises the diverse ways in which an evidential report can be affected by a preconceived view of what is the 'good' answer. That might come from researchers performing data mining until they get the answer they want or from them simply falsifying their results because of their sponsor's interests. It can also come from institutional pressures in science: peer review systematically favoring some sort of result or deeply rooted hypotheses making scientists revise their experimental procedure until the output fits 'what is known'.

Another class of cases has to do with the risks of using something as a stand-in (as a model) in order to learn about something else. If one uses, for instance, an animal to learn about the potential side effects of a drug on humans, the extrapolation might go wrong because there is some biological mechanism in the model not present in the target (or the other way around), which makes the drug have some effect in one group of subjects but not in the other. The result from the model subjects will thus be systematically biased when used as a report for the target subjects.

To model sources that are potentially systematically biased, I redefine the reliability variable:

- The new reliability variable  $R_i = \{r_i, b_i^h, b_i^{-h}\}$ , where  $r_i$  stands, as before, for the proposition that the source is reliable,  $b_i^h$  stands for the

proposition that the source is biased toward a positive report for the hypothesis regardless of its truth, and  $b_i^{-h}$  stands for the proposition that the source is biased toward a negative report.

This ternary variable (all the previous variables were binary) is arrived at by giving a more finely grained specification of the proposition  $\neg r_i$ . It is now decomposed into two disjoint propositions—that is, we now have  $\neg r_i = b_i^h \cup b_i^{-h}$ .

My first modification inserts this new variable into the previous model. The probabilistic independencies that can be read off the Bayesian networks in figure 1 still hold. Furthermore, the specifications of the prior probabilities  $h_0$  and  $\rho_i$  in condition (1) are retained. We need, however, to specify more probabilities for  $R_i$ :

$$P(b_i^h|\neg r_i) = \alpha_i \quad \text{and} \quad P(b_i^{-h}|\neg r_i) = \bar{\alpha}_i. \tag{9}$$

This condition assigns prior probabilities to the propositions about positive and negative biases given that the source is already known to be unreliable. Note that what was interpreted as a ‘randomization parameter’ by Bovens and Hartmann is used explicitly as a strength of belief here. Combining condition (9) with condition (1), we have what follows: the prior probability of a positive bias  $P(b_i^h)$  is  $\alpha_i \bar{\rho}_i$ , and the prior probability of a negative bias  $P(b_i^{-h})$  is  $\bar{\alpha}_i \bar{\rho}_i$ .

Finally, we need to expand table 1 by stating explicitly how likely  $e_i$  is, conditional on  $b_i^h$  and  $b_i^{-h}$ . This expansion gives us table 2. Since the uncertainty that figured initially in  $P(e_i|H, R_i)$  has been shifted to  $R_i$ , the evidential variable  $E_i$  is now a deterministic function of  $H$  and  $R_i$ . This deterministic relation might come as a surprise to some, but it should not be surprising. If we remain committed to the irrelevance of an unreliable source (IUS), the columns for  $b_i^h$  and for  $b_i^{-h}$  in table 2 must each contain the same value twice. If instead of having 1 and 0 for these values, we opt for values strictly between these two, we reintroduce into the model Bovens and Hartmann’s unreliability as randomizing. The counterexamples used in section 3 would thus apply. As long as we remain committed to the IUS condition, the notion of a systematic bias must be captured by a deterministic function. In future work, I will drop the IUS condition in the context of calibration issues, but I keep it here since it seems pertinent for some sources of unreliability.

TABLE 2. PROBABILITY OF A POSITIVE REPORT GIVEN THE VALUES OF  $H$  AND  $R_i$ ; EXPANDED

$P(e_i H, R_i)$	$r_i$	$b_i^h$	$b_i^{-h}$
$h$	1	1	0
$\neg h$	0	1	0

The probabilities in table 2 and in equations (9) give us two new versions of the model. Define  $P_I(\cdot)$  as the joint probability distribution associated with the independent-reliability situation (i.e., the distribution associated with fig. 1A) and  $P_{S'}(\cdot)$  as the distribution associated with the shared-reliability situation (i.e., the distribution associated with fig. 1B). We can now assess the variety-of-evidence thesis: Is it always the case that  $P_I^*(h) > P_{S'}^*(h)$ ?

In fact,  $P_I^*(h)$  is no different from  $P_I^*(h)$ —that is, for the case of sources with independent reliabilities, Bovens and Hartmann's version and my version give the same result (proof in app. B). This is welcome news, given that Bovens and Hartmann's version seems to capture what one means in saying that two evidential reports are fully independent regarding a hypothesis—that is, it concurs with what Shogenji (2005, 308) presents as “a general consensus among probability theorists on how to formalize the condition that two pieces of evidence  $E_1$  and  $E_2$  are independent of each other with respect to proposition  $A$ .”<sup>8</sup>

Things are different when we turn to the new version with shared reliability. The posterior probability of the hypothesis is now (proof in app. B)

$$P_{S'}^*(h) = \frac{h_0}{h_0 + \bar{h}_0 L_{S'}}, \quad \text{where} \quad L_{S'} = \frac{\alpha \bar{\rho}}{\rho + \alpha \bar{\rho}}. \quad (10)$$

The likelihood ratio  $L_{S'}$  is identical to the one resulting from an evidential set with only a single element instead of two (see eq. [A5]). In other words, adding a second positive report in this new shared-reliability model has no effect on the degree of confirmation of the hypothesis. The reason for this result is simple: the second report cannot be anything but consistent with the first report in this model. The two evidential variables not only share reliability; they also share the direction of the bias if they are indeed biased. There is no longer the possibility of detecting that a source is unreliable by finding discordant reports coming from this source. Since this possibility no longer obtains, multiplying the reports from the same source becomes useless.

Is it still possible that the reports in the shared-reliability situation are more confirmatory than the ones in the independent-reliability situation? No. The posterior probability of  $h$  is strictly higher in the independency case if the probability that the sources are reliable is higher than 0 (see app. B for the proof). There is thus no combination of admissible parameter values for which having shared reliability is, *ceteris paribus*, better. Using the same source again is not conducive to confirmation because it no longer holds the promise of detecting the potential unreliability of the source. A second

8. It is a case of Sober's conjunctive fork (Sober 1989; Fitelson 2001).

independent report is thus necessarily more confirmatory. The variety-of-evidence thesis holds without qualification in this version of the models.

**5. Degrees of Independence.** The result supporting the variety-of-evidence thesis in the previous section suffers from a major limitation. While the variety-of-evidence thesis explicitly compares more independent to less independent evidential elements, the comparison made with our two models is between fully independent and fully dependent evidential elements. Our comparison of confirmation was restricted to the two ends of a spectrum, whereas the variety-of-evidence thesis deals with how confirmation changes with changes in the degree of independence.

There is a simple way to model degrees of independence by extending the setup of the previous section. The graphical representation of this extended model is in figure 3, and its associated probability distribution will be labeled  $P_F(\cdot)$ . The modification here adds a probabilistic association between the two reliability variables  $R_1$  and  $R_2$ .<sup>9</sup> The rest of the model remains intact.

The association between the reliability variables is fully captured by specifying the probabilities for the nine possible combinations of their values. Panel A of table 3 offers a general notation for these nine possibilities. For instance,  $\omega_{rr}$  is the probability that both sources are reliable. The elements on the main diagonal ( $\omega_{rr}$ ,  $\omega_{hh}$ ,  $\omega_{-h-h}$ ) are the probabilities associated with the proposition that the two sources are in the same reliability state. Note that the table already assumes symmetry between the two sources—that is,  $P(r_1, b_2^h) = P(b_1^h, r_2) = \omega_{rh}$ , and so forth. This assumption was also used in the previous sections as part of the assumptions sufficient to meet the ceteris paribus condition of the variety-of-evidence thesis.

In this new model, the posterior belief in the hypothesis given two positive reports is (proof in app. C.1)

$$P_F^*(h) = \frac{h_0}{h_0 + \bar{h}_0 L_F}, \quad \text{where} \quad L_F = \frac{\omega_{hh}}{\omega_{rr} + 2\omega_{rh} + \omega_{hh}}. \quad (11)$$

Table 3 panels B and C give the specific values taken by the  $\omega$ 's for the two extreme cases on which the previous sections focused. It can be easily verified using these values that the expression in equation (11) reduces to equation (10) or (6) for each of these extreme cases—that is, the model of the

9. Bovens and Hartmann (2003, 75–77) offer a model with a super-reliability variable that is specified as a common cause of the  $R_i$ 's. However, they do not use it to discuss the variety-of-evidence thesis. Since this super-reliability variable is difficult to interpret and since only modeling a probabilistic association between  $R_1$  and  $R_2$  is sufficient for my goal here, I opt for the second option.

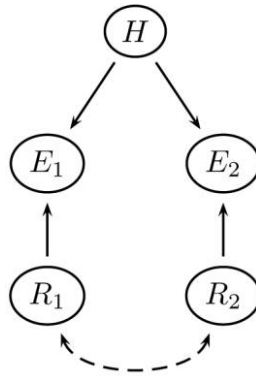


Figure 3. Extended model with degrees of independence.

previous section is fully embedded into this one (including its result for the variety-of-evidence thesis).

Is there a ready measure of degrees of independence? My proposal is based on the following consideration. Compare table 3 panels B and C. The probability mass is all on the main diagonal in the first case. In other words, it never happens that the two sources are in different reliability states. In the case of full independence, the probability mass is more spread out since the joint probability  $P(R_1, R_2)$  is simply the product of the marginal probabilities (i.e.,  $P(R_1)P(R_2)$ ). In fact, each element on the main diagonal in table 3, panel C, is exactly the square of the same element in table 3, panel B. This fact suggests a specific metric to characterize degrees of independence.

TABLE 3. JOINT PROBABILITIES FOR THE RELIABILITY VARIABLES (ASSUMING SYMMETRY)

$P(R_1, R_2)$	$r_2$	$b_2^h$	$b_2^{-h}$
A. General Case			
$r_1$	$\omega_{rr}$	$\omega_{rh}$	$\omega_{r-h}$
$b_1^h$	$\omega_{rh}$	$\omega_{hh}$	$\omega_{h-h}$
$b_1^{-h}$	$\omega_{r-h}$	$\omega_{h-h}$	$\omega_{-h-h}$
B. Fully Shared Reliability			
$r_1$	$\rho$	0	0
$b_1^h$	0	$\bar{\rho}\alpha$	0
$b_1^{-h}$	0	0	$\bar{\rho}\bar{\alpha}$
C. Fully Independent Reliability			
$r_1$	$\rho^2$	$\rho\bar{\rho}\alpha$	$\rho\bar{\rho}\bar{\alpha}$
$b_1^h$	$\rho\bar{\rho}\alpha$	$(\bar{\rho}\alpha)^2$	$\bar{\rho}^2\alpha\bar{\alpha}$
$b_1^{-h}$	$\rho\bar{\rho}\bar{\alpha}$	$\bar{\rho}^2\alpha\bar{\alpha}$	$(\bar{\rho}\bar{\alpha})^2$



Define a variable  $\delta \in [0, 1]$  that is interpreted as measuring the distance of the evidential set from fully shared reliability—that is, when  $\delta = 0$  we have no independence, when  $\delta = 1$  we have full independence, and when  $\delta$  is strictly between 0 and 1, we have only partial independence. Given values for  $\rho$ ,  $\alpha$ , and  $\delta$ , the elements on the main diagonal are

$$\omega_{rr} = \rho^{1+\delta}, \quad \omega_{hh} = (\bar{\rho}\alpha)^{1+\delta}, \quad \omega_{-h-h} = (\bar{\rho}\bar{\alpha})^{1+\delta}. \quad (12)$$

These relations entail that the probability mass is shifted away from the elements on the main diagonal as the degree of independence increases. In other words, it becomes less likely that the two sources share the same reliability state.

With this variable  $\delta$ , the variety-of-evidence thesis can be restated.

**Variety-of-evidence thesis.** *Ceteris paribus*,  $\partial P_F^*(h)/\partial \delta > 0$ , for all admissible values of  $\rho$ ,  $\alpha$ , and  $\delta$ .

The restatement of the thesis is thus that the posterior degree of belief in the hypothesis invariably increases as we marginally increase the degree of independence of the evidential sources.

Before we assess this thesis, we need to specify how the off-diagonal elements in table 3, panel A, change as  $\delta$  is modified. One obvious restriction is that the sum of all the elements (the nine  $\omega$ 's) must be 1. The interpretation of the *ceteris paribus* condition previously used also restricts the values of the off-diagonal elements but not enough to ensure uniqueness. In addition to these restrictions, I thus also stipulate that the marginal probabilities of  $R_1$  and  $R_2$  are not a function of  $\delta$ —that is,  $P(r_i) = \rho$ ,  $P(b_i^h) = \bar{\rho}\alpha$ , and  $P(b_i^{-h}) = \bar{\rho}\bar{\alpha}$ , for  $i = \{1, 2\}$  and for all  $\delta \in [0, 1]$  (see app. C.2).

This model leads to a qualification of the variety-of-evidence thesis (proof in app. C.3). Increasing the degree of independence leads to more confirmation if and only if the following condition holds:

$$(1 - 2\bar{\rho}\bar{\alpha}) \ln(\bar{\rho}\alpha) + (\bar{\rho}\bar{\alpha})^{1+\delta} \ln\left(\frac{\alpha}{\bar{\alpha}}\right) < 0. \quad (13)$$

But there are combinations of admissible values for  $\rho$ ,  $\alpha$ , and  $\delta$  that violate this condition.

Figure 4 presents graphically the different possibilities. For most combinations of  $\rho$  and  $\alpha$ , the relationship between degree of independence and confirmation is as stated by the variety-of-evidence thesis (fig. 4B presents such a case).<sup>10</sup> Figure 4A shows that there are in fact two distinct regions of the parameter space where the relationship between independence and con-

10. The proportion of the parameter space  $\alpha \times \rho$  where the relationship between independence and confirmation is not monotonically increasing—i.e., the area of the two

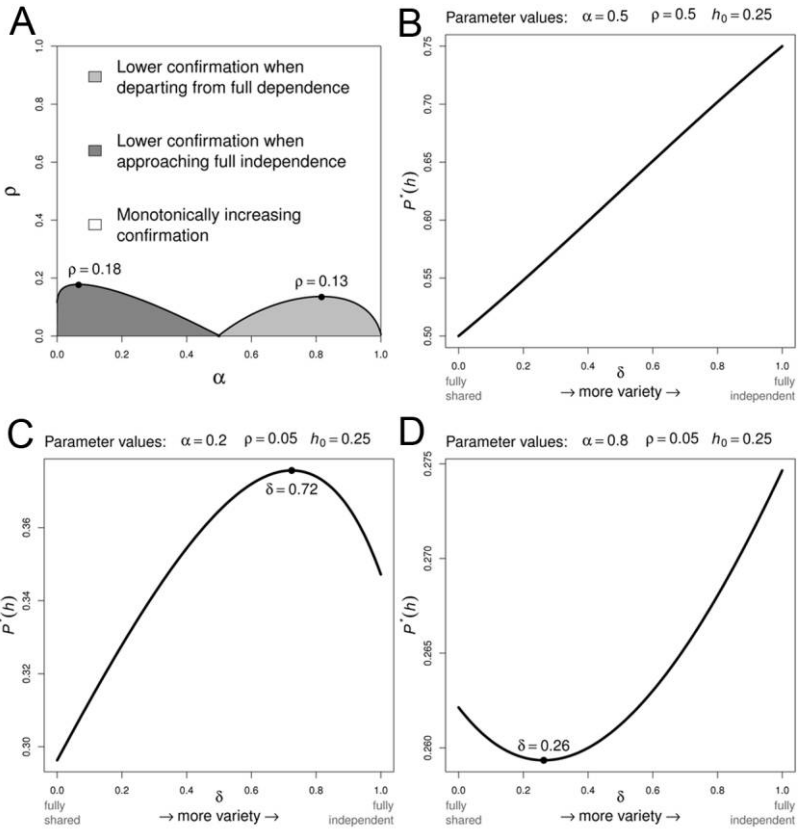


Figure 4. Nonmonotonicity is possible. *A*, Parameter combinations resulting in a nonmonotonic relationship between degrees of independence and confirmation; *B*, monotonic relationship; *C*, low  $\alpha$ ; *D*, high  $\alpha$ .

firmation is nonmonotonic. These two possibilities share (extremely) low values of  $\rho$ . In other words, these are situations in which prior information leads the agent to believe that it is highly unlikely that a given source is truth tracking. The relationship is indeed always monotonic when the trust in the source is above .18 (in Bovens and Hartmann’s model this was .5).

There are two features distinguishing the two nonmonotonic situations from each other. First, as shown in figure 4*A*, they differ in their values for  $\alpha$ —that is, the probability that the report is positive given that the source is

gray regions in fig. 4*A*—is 10.3%. As a point of comparison, this proportion is 15.3% in Bovens and Hartmann’s model (as depicted in fig. 2). If one considers instead the three-dimensional space  $\alpha \times \rho \times \delta$ , only 2% of it gives  $\partial P^*(h)/\partial \delta < 0$ . These proportions should not be interpreted as probabilities.

unreliable. Second, as is evident by comparing figure 4C and 4D, the two regions are associated with different shapes of nonmonotonicity (concave vs. convex functions).

What is going on? As in Bovens and Hartmann's model, what happens is that, upon learning  $e_1$  and  $e_2$ , the agent reassesses the probability that the sources are reliable. The region of the space  $\alpha \times \rho \times \delta$  where confirmation decreases with independence is exactly the region where trust in the reliability of the sources decreases with independence (proof in app. C.4). In other words, to compare two evidential sets (for  $h$ )—say  $\mathbf{E} = \{e_1, e_2\}$  and  $\mathbf{E}' = \{e'_1, e'_2\}$ , which differ only with respect to their degree of reliability independence ( $\delta$ ) (the elements of  $\mathbf{E}$  being more independent than the elements of  $\mathbf{E}'$ )—one simply needs to assess the following ratio for each set:

$$\frac{\omega_{rr} + 2\omega_{rh}}{\omega_{hh}}. \quad (14)$$

The set with the higher ratio is more confirmatory than the other. The numerator of this ratio captures the probability of realizations of  $R_1$  and  $R_2$  that generate two evidential elements ( $e_1$  and  $e_2$ ) that are indeed truth revealing for  $h$ . The denominator is the probability that the two sources are producing positive-but-garbage reports for  $h$ . The denominator of  $\mathbf{E}$  will always be smaller than that of  $\mathbf{E}'$ . This fact might capture the intuitive appeal of the variety-of-evidence thesis: it is less likely to get two garbage reports from sources that are (more) independent. But the full ratio is what ultimately decides between  $\mathbf{E}$  and  $\mathbf{E}'$ .

Let me briefly discuss the only two situations in which the variety-of-evidence thesis is turned upside down. First, for low values of  $\alpha$  combined with extremely low values of  $\rho$  (as in fig. 4C), getting two positive reports comes as a surprise—it was judged far more likely to receive at least one negative report because of the realization of  $b_i^{-h}$ . In this case, moving toward independence is initially beneficial, but more independence becomes detrimental to confirmation as one approaches the extreme of full independence. This result is interesting because it means that slightly departing from full independence sometimes increases confirmation.

Second, for high values of  $\alpha$  combined with extremely low values for  $\rho$  (as in fig. 4D), getting two positive reports is not surprising; however, the agent judges it highly likely that the information is worthless (because  $b_1^h$  and  $b_2^h$  are likely to be realized). In this case, a departure from full dependence adversely affects confirmation. An implication of this nonmonotonicity is that the second positive report can be disconfirming  $h$  (i.e.,  $P_F(h|e_1, e_2) < P_F(h|e_1)$ ). Remember from section 4 that the posterior belief in  $h$  after two fully dependent reports (i.e.,  $\delta = 0$ ) is identical to the posterior belief after a single report. Both are represented by the point at the extreme left of the curve in figure 4D. All the points lying below this point are thus

cases in which the second report is disconfirming  $h$ . The agent puts so little trust in the evidential sources that a second report is interpreted as a sign that both sources are positively biased, and the initial (slight) increase in the belief for  $h$  is cut back.

**6. Conclusion.** The variety-of-evidence thesis seems to be a widespread implicit guideline in scientific practice. This thesis says that, *ceteris paribus*, the confirmatory power of an evidential set for a given hypothesis increases with the diversity (i.e., the independence) of the evidential elements in the set. Thus, one should praise ‘independent evidence’ and be suspicious of the rest.

Bovens and Hartmann (2002, 2003) cast doubt on the universal applicability of this thesis by showing with a simple model that, in some peculiar epistemic situations, it is sometimes a disadvantage for confirmation to have independent evidential elements, *ceteris paribus*. I have argued that the relevance of this result is diminished by two characteristics of their model.

First, their idea that unreliable sources are randomizers leads them to model fully dependent sources in a way that is unlikely to reflect how scientists think about their sources of evidence. The problem is that Bovens and Hartmann assume that two fully dependent sources still produce two independent reports when they are unreliable (i.e.,  $E_1 \perp E_2 | \neg r$ ). Instead, in actual scientific settings it seems to be the case that two reports coming from fully dependent sources will always coincide, even when the sources are unreliable. In section 4, I showed that the variety-of-evidence thesis is rehabilitated once the independent-randomizer assumption is dropped and replaced with the assumption that an unreliable source is systematically biased. This modification is compatible with the key intuition behind the notion of reliability in Bovens and Hartmann’s model (i.e., IUS).

Second, there is another serious limitation in Bovens and Hartmann’s model, a limitation that my first modification of their model shares. The comparison made to assess the variety-of-evidence thesis is between extremes; it is between fully independent and fully dependent evidential elements. The most relevant comparison is rather one of degree: less versus more independence of the sources. In section 5, I showed that when the model is modified to enable comparisons of degrees of independence, the variety-of-evidence thesis needs to be qualified. There are special epistemic situations wherein more independence does not give more confirmation. This qualification only applies to a subinterval of the spectrum from full dependence to full independence. Indeed, the two extremes of the spectrum always stand in the confirmatory relationship depicted by the variety-of-evidence thesis.

Where do my modeling efforts leave us? First, the usual caveat about idealization applies: it might well be that the way in which epistemic sit-

uations have been modeled here does not capture what is pertinent for the variety-of-evidence thesis. It is certain that my model does not encompass all the ways in which an evidential source can be unreliable (e.g., the calibration problems mentioned in sec. 2).

Even if one accepts the idealizations, the conclusion to draw about the variety-of-evidence thesis is not straightforward. One plausible reaction to the result of the last section is as follows. The variety-of-evidence thesis can break down in the extended model only if the agent has enormous doubts about the reliability of the evidential source; she must judge it to be at least 82% likely that the source is unreliable. One could thus read the result as highlighting the danger of using extremely weak evidential sources, rather than as a direct refutation of the variety-of-evidence thesis. This thesis could be interpreted as implicitly assuming that the evidential sources are sufficiently trustworthy to begin with. The fate of the variety-of-evidence thesis is not yet settled.

**Appendix A: Bovens and Hartmann’s Version**

One gets the two likelihood ratios  $L_I$  and  $L_S$  by using the probabilistic information encoded in figure 1 together with equations (1), (2), and (4). I start with the likelihood ratio for the independent-reliability version:

$$L_I = \frac{P_I(e_1, e_2 | \neg h)}{P_I(e_1, e_2 | h)} = \frac{\sum_{R_1, R_2} P_I(e_1 | \neg h, R_1) P_I(R_1) P_I(e_2 | \neg h, R_2) P_I(R_2)}{\sum_{R_1, R_2} P_I(e_1 | h, R_1) P_I(R_1) P_I(e_2 | h, R_2) P_I(R_2)}$$

Given that the terms in the multiplications are either solely about source 1 or source 2, I factorize by source:

$$\begin{aligned} &= \frac{\prod_{i=\{1,2\}} \sum_{R_i} P_I(e_i | \neg h, R_i) P_I(R_i)}{\prod_{i=\{1,2\}} \sum_{R_i} P_I(e_i | h, R_i) P_I(R_i)} \\ &= \frac{\prod_i [P_I(e_i | \neg h, r_i) P_I(r_i) + P_I(e_i | \neg h, \neg r_i) P_I(\neg r_i)]}{\prod_i [P_I(e_i | h, r_i) P_I(r_i) + P_I(e_i | h, \neg r_i) P_I(\neg r_i)]} \\ &= \frac{\prod_i [0\rho_i + \alpha_i \bar{\rho}_i]}{\prod_i [1\rho_i + \alpha_i \bar{\rho}_i]} = \frac{\alpha_1 \alpha_2 \bar{\rho}_1 \bar{\rho}_2}{(\rho_1 + \alpha_1 \bar{\rho}_1)(\rho_2 + \alpha_2 \bar{\rho}_2)} \tag{A1} \end{aligned}$$

$$L_I = \frac{(\alpha \bar{\rho})^2}{(\rho + \alpha \bar{\rho})^2} \tag{A2}$$

The third line results from plugging in the parameter values and then simplifying; the fourth line imposes the ceteris paribus condition.

I do the same for the shared-reliability version:

$$\begin{aligned}
 L_S &= \frac{P_S(e_1, e_2 | \neg h)}{P_S(e_1, e_2 | h)} \\
 &= \frac{\sum_R P_S(e_1 | \neg h, R) P_S(e_2 | \neg h, R) P_S(R)}{\sum_R P_S(e_1 | h, R) P_S(e_2 | h, R) P_S(R)} \\
 &= \frac{P_S(e_1 | \neg h, r) P_S(e_2 | \neg h, r) P_S(r) + P_S(e_1 | \neg h, \neg r) P_S(e_2 | \neg h, \neg r) P_S(\neg r)}{P_S(e_1 | h, r) P_S(e_2 | h, r) P_S(r) + P_S(e_1 | h, \neg r) P_S(e_2 | h, \neg r) P_S(\neg r)} \\
 &= \frac{0 \times 0\rho + \alpha_1 \alpha_2 \bar{\rho}}{1 \times 1\rho + \alpha_1 \alpha_2 \bar{\rho}} = \frac{\alpha_1 \alpha_2 \bar{\rho}}{\rho + \alpha_1 \alpha_2 \bar{\rho}} \tag{A3}
 \end{aligned}$$

$$L_S = \frac{\alpha^2 \bar{\rho}}{\rho + \alpha^2 \bar{\rho}}. \tag{A4}$$

Note that from (A1) or (A3), it can easily be seen that the posterior belief in  $h$  when only one positive report is known is

$$P(h|e_i) = \frac{h_0}{h_0 + \bar{h}_0 L^i}, \quad \text{where } L^i = \frac{\alpha_i \bar{\rho}_i}{\rho_i + \alpha_i \bar{\rho}_i}. \tag{A5}$$

The relation between  $P_I^*(h)$  and  $P_S^*(h)$  holds as stated by the variety-of-evidence thesis if and only if

$$\begin{aligned}
 \frac{h_0}{h_0 + \bar{h}_0 L_I} &> \frac{h_0}{h_0 + \bar{h}_0 L_S} \Leftrightarrow L_S > L_I \Leftrightarrow \frac{\alpha^2 \bar{\rho}}{\rho + \alpha^2 \bar{\rho}} > \frac{\alpha^2 \bar{\rho}^2}{(\rho + \alpha \bar{\rho})^2} \\
 &\Leftrightarrow (\rho + \alpha \bar{\rho})^2 > \bar{\rho}(\rho + \alpha^2 \bar{\rho}) \Leftrightarrow \rho^2 + 2\alpha \bar{\rho} \rho + \alpha^2 \bar{\rho}^2 > \bar{\rho} \rho + \alpha^2 \bar{\rho}^2 \\
 &\Leftrightarrow (1 - \bar{\rho}) + 2\alpha \bar{\rho} > \bar{\rho} \Leftrightarrow 1 > 2\bar{\rho} - 2\alpha \bar{\rho} \Leftrightarrow .5 > \alpha \bar{\rho}.
 \end{aligned}$$

**Appendix B: Model with Unreliability as Systematic Bias**

I compute the likelihood ratio for the independent-reliability version:

$$\begin{aligned}
 L_I &= \frac{P_I(e_1, e_2 | \neg h)}{P_I(e_1, e_2 | h)} \\
 &= \frac{\prod_i [P_I(e_i | \neg h, r_i) P_I(r_i) + P_I(e_i | \neg h, b_i^h) P_I(b_i^h) + P_I(e_i | \neg h, b_i^{-h}) P_I(b_i^{-h})]}{\prod_i [P_I(e_i | h, r_i) P_I(r_i) + P_I(e_i | h, b_i^h) P_I(b_i^h) + P_I(e_i | h, b_i^{-h}) P_I(b_i^{-h})]} \\
 &= \frac{\prod_i [0\rho_i + 1\alpha_i \bar{\rho}_i + 0\bar{\alpha}_i \bar{\rho}_i]}{\prod_i [1\rho_i + 1\alpha_i \bar{\rho}_i + 0\bar{\alpha}_i \bar{\rho}_i]} = \frac{\alpha_1 \alpha_2 \bar{\rho}_1 \bar{\rho}_2}{(\rho_1 + \alpha_1 \bar{\rho}_1)(\rho_2 + \alpha_2 \bar{\rho}_2)}.
 \end{aligned}$$

The last expression is identical to the right-hand side of equation (A1), which proves that my model and Bovens and Hartmann’s model agree when sources are reliability independent.

I do the same for the shared-reliability version:

$$\begin{aligned} L_{S'} &= \frac{P_{S'}(e_1, e_2 | \neg h)}{P_{S'}(e_1, e_2 | h)} \\ &= \frac{\sum_R P_{S'}(e_1 | \neg h, R) P_{S'}(e_2 | \neg h, R) P_{S'}(R)}{\sum_R P_{S'}(e_1 | h, R) P_{S'}(e_2 | h, R) P_{S'}(R)} \\ &= \frac{0 \times 0\rho + 1 \times 1\alpha\bar{\rho} + 0 \times 0\alpha\bar{\rho}}{1 \times 1\rho + 1 \times 1\alpha\bar{\rho} + 0 \times 0\alpha\bar{\rho}} = \frac{\alpha\bar{\rho}}{\rho + \alpha\bar{\rho}}. \end{aligned}$$

The last expression does not give equation (A3)—that is, this version of the shared-reliability situation does not concord with Bovens and Hartmann’s version. In fact, it is equal to equation (A5), which expresses the likelihood ratio for a single positive report.

I now prove that independent reliability is always better. To fulfill the *ceteris paribus* clause, I again assume that the prior  $h_0$  is the same for both models, that  $\alpha_1 = \alpha_2 = \alpha$ , and that  $\rho_1 = \rho_2 = \rho$ . Then we have

$$\begin{aligned} P_{I'}^*(h) > P_{S'}^*(h) &\Leftrightarrow L_{I'} < L_{S'} \Leftrightarrow \frac{\alpha^2 \bar{\rho}^2}{(\rho + \alpha\bar{\rho})^2} < \frac{\alpha\bar{\rho}}{\rho + \alpha\bar{\rho}} \\ &\Leftrightarrow \alpha\bar{\rho} < \rho + \alpha\bar{\rho} \Leftrightarrow 0 < \rho. \end{aligned}$$

Thus, as soon as we have a nonnull prior probability that the evidential sources are reliable, reliability-independent sources are epistemically preferable.

### Appendix C: Extended Model

**1. Posterior Belief in the Hypothesis.** I focus on the likelihood ratio:

$$\begin{aligned} L_F &= \frac{P_F(e_1, e_2 | \neg h)}{P_F(e_1, e_2 | h)} = \frac{\sum_{R_1, R_2} P_F(e_1, e_2, R_1, R_2 | \neg h)}{\sum_{R_1, R_2} P_F(e_1, e_2, R_1, R_2 | h)} \\ &= \frac{\sum_{R_1, R_2} P_F(e_1 | \neg h, R_1) P_F(e_2 | \neg h, R_2) P_F(R_1, R_2)}{\sum_{R_1, R_2} P_F(e_1 | h, R_1) P_F(e_2 | h, R_2) P_F(R_1, R_2)} \\ &= \frac{P_F(b_i^h, b_j^h)}{P_F(r_i, r_j) + 2P_F(r_i, b_j^h) + P_F(b_i^h, b_j^h)} \\ &= \frac{\omega_{hh}}{\omega_{rr} + 2\omega_{rh} + \omega_{hh}} = \left[ 1 + \frac{\omega_{rr} + 2\omega_{rh}}{\omega_{hh}} \right]^{-1}, \end{aligned}$$

where the second-to-last line uses the information in table 2 and the last line uses table 3, panel A.

**2. Conditions for the Off-Diagonal Elements.** Using the notation from table 3, panel A, I rewrite my condition that the marginal probabilities of  $R_1$  and  $R_2$  are not a function of  $\delta$ :

$$\begin{aligned} P(r_i) &= \rho = \omega_{rr} + \omega_{rh} + \omega_{r-h}, \\ P(b_i^h) &= \bar{\rho}\alpha = \omega_{rh} + \omega_{hh} + \omega_{h-h}, \\ P(b_i^{-h}) &= \bar{\rho}\bar{\alpha} = \omega_{r-h} + \omega_{-hh} + \omega_{-h-h}. \end{aligned}$$

I then solve this system of equation for the off-diagonal elements in terms of the diagonal elements,  $\rho$  and  $\alpha$  (I omit the simple algebraic manipulations).

$$\begin{aligned} \omega_{rh} &= \rho + \bar{\rho}\alpha - .5(1 + \omega_{rr} + \omega_{hh} - \omega_{-h-h}), \\ \omega_{r-h} &= \rho + \bar{\rho}\bar{\alpha} - .5(1 + \omega_{rr} - \omega_{hh} + \omega_{-h-h}), \\ \omega_{h-h} &= .5(1 + \omega_{rr} - \omega_{hh} - \omega_{-h-h}) - \rho. \end{aligned} \tag{C1}$$

**3. The Derivative of the Likelihood Ratio.** We can rewrite the likelihood ratio in (11) by using information from the system of equations (C1):

$$\begin{aligned} L_F &= \frac{\omega_{hh}}{\omega_{rr} + 2\omega_{rh} + \omega_{hh}} = \frac{\omega_{hh}}{\omega_{rr} + \omega_{hh} + 2(1 - \bar{\rho}\bar{\alpha}) - 1 - \omega_{rr} - \omega_{hh} + \omega_{-h-h}} \\ &= \frac{\omega_{hh}}{1 - 2\bar{\rho}\bar{\alpha} + \omega_{-h-h}} = \frac{(\bar{\rho}\alpha)^{1+\delta}}{1 - 2\bar{\rho}\bar{\alpha} + (\bar{\rho}\bar{\alpha})^{1+\delta}}, \end{aligned}$$

where the last equality uses condition (12). I take the derivative with respect to  $\delta$ :

$$\begin{aligned} \frac{\partial L_F}{\partial \delta} &= \frac{(\bar{\rho}\alpha)^{1+\delta} \ln(\bar{\rho}\alpha)(1 - 2\bar{\rho}\bar{\alpha} + (\bar{\rho}\bar{\alpha})^{1+\delta}) - (\bar{\rho}\alpha)^{1+\delta} (\bar{\rho}\bar{\alpha})^{1+\delta} \ln(\bar{\rho}\bar{\alpha})}{[1 - 2\bar{\rho}\bar{\alpha} + (\bar{\rho}\bar{\alpha})^{1+\delta}]^2} \\ &= \frac{(\bar{\rho}\alpha)^{1+\delta} [(1 - 2\bar{\rho}\bar{\alpha}) \ln(\bar{\rho}\alpha) + (\bar{\rho}\bar{\alpha})^{1+\delta} (\ln(\bar{\rho}\alpha) - \ln(\bar{\rho}\bar{\alpha}))]}{[1 - 2\bar{\rho}\bar{\alpha} + (\bar{\rho}\bar{\alpha})^{1+\delta}]^2} \\ &= \frac{(\bar{\rho}\alpha)^{1+\delta} [(1 - 2\bar{\rho}\bar{\alpha}) \ln(\bar{\rho}\alpha) + (\bar{\rho}\bar{\alpha})^{1+\delta} \ln(\alpha/\bar{\alpha})]}{[1 - 2\bar{\rho}\bar{\alpha} + (\bar{\rho}\bar{\alpha})^{1+\delta}]^2}. \end{aligned}$$

The variety-of-evidence thesis maintains that  $\partial L_F / \partial \delta < 0$  for all admissible parameter values. Verifying this:

$$\frac{\partial L_F}{\partial \delta} < 0 \Leftrightarrow (1 - 2\bar{\rho}\bar{\alpha}) \underbrace{\ln(\bar{\rho}\alpha)}_{<0} + \underbrace{(\bar{\rho}\bar{\alpha})^{1+\delta} \ln\left(\frac{\alpha}{\bar{\alpha}}\right)}_{>0} < 0, \tag{C2}$$



which does not hold for some combination of parameter values (see fig. 4). The fact that two distinct regions of  $\alpha \times \rho$  lead to a reversal of the inequality can be seen from expression (C2). The first term is positive (i.e., contributing to a reversal of the relationship) if and only if  $(1 - 2\bar{\rho}\bar{\alpha}) < 0$ , or more intuitively,  $.5 < \bar{\rho}\bar{\alpha}$ . The second term is positive if and only if  $\alpha > .5$ . It follows that the two terms cannot be positive at the same time.

**4. Posterior Belief in Reliability.** Having a single reliable source is sufficient for the two positive reports  $e_1$  and  $e_2$  to be truth revealing. The posterior belief that at least one source is reliable is

$$\begin{aligned} P(r_1 \cup r_2 | e_1, e_2) &= P(r_1, r_2 | e_1, e_2) + 2P(r_i, b_j^h | e_1, e_2) + \underbrace{2P(r_i, b_j^{-h} | e_1, e_2)}_{=0} \\ &= \frac{P(e_1, e_2 | r_1, r_2)P(r_1, r_2) + 2P(e_1, e_2 | r_i, b_j^h)P(r_i, b_j^h)}{P(e_1, e_2)} \\ P^*(r_1 \cup r_2) &= \frac{h_0(\omega_{rr} + 2\omega_{rh})}{h_0(\omega_{rr} + 2\omega_{rh}) + \omega_{hh}} = \left[ 1 + \frac{\omega_{hh}}{h_0(\omega_{rr} + 2\omega_{rh})} \right]^{-1}. \end{aligned}$$

For ease of manipulation, I use the last equality to define the variable DT (for distrust):

$$DT = P^*(r_1 \cup r_2)^{-1} - 1 = \frac{\omega_{hh}}{h_0(\omega_{rr} + 2\omega_{rh})}.$$

Reusing a result in appendix C.1, I also define a variable  $C$  (for strength of confirmation):

$$C = L_F^{-1} - 1 = \frac{\omega_{rr} + 2\omega_{rh}}{\omega_{hh}}.$$

Distrust and confirmation are related as

$$C = \frac{1}{h_0 DT}$$

from which it follows that

$$\frac{\partial C}{\partial \delta} > 0 \Leftrightarrow \frac{\partial DT}{\partial \delta} < 0 \Leftrightarrow \frac{\partial P^*(r_1 \cup r_2)}{\partial \delta} > 0.$$

In words, confirmation increases with reliability independence if and only if posterior trust in the sources increases.

## REFERENCES

- Bovens, Luc, and Stephan Hartmann. 2002. "Bayesian Networks and the Problem of Unreliable Instruments." *Philosophy of Science* 69 (1): 29–72.
- . 2003. *Bayesian Epistemology*. Oxford: Oxford University Press.
- CERN. 2012. "Neutrinos Sent from CERN to Gran Sasso Respect the Cosmic Speed Limit." *ScienceDaily*, June 8. <http://www.sciencedaily.com/releases/2012/06/120608152339.htm>.
- Contaldi, Carlo R. 2011. "The OPERA Neutrino Velocity Result and the Synchronisation of Clocks." ArXiv, Cornell University Library. <http://arxiv.org/abs/1109.6160>.
- Earman, John. 1992. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.
- Fitelson, Branden. 1996. "Wayne, Horwich, and Evidential Diversity." *Philosophy of Science* 63 (4): 652–60.
- . 2001. "A Bayesian Account of Independent Evidence with Applications." *Philosophy of Science* 68 (3): S123–S140.
- Hartmann, Stephan. 2008. "Modeling in Philosophy of Science." In *Representation, Evidence, and Justification: Themes from Suppes*, ed. Michael Frauchiger and Wilhelm K. Essler, 95–121. Heusenstamm: Ontos.
- Horwich, Paul. 1982. *Probability and Evidence*. Cambridge: Cambridge University Press.
- . 1998. "Wittgensteinian Bayesianism." In *Philosophy of Science: The Central Issues*, ed. Martin Curd and Jan A. Cover, 607–24. New York: Norton.
- Howson, Colin, and Peter Urbach. 1993. *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Chicago: Open Court.
- ICARUS Collaboration. 2012. "Measurement of the Neutrino Velocity with the ICARUS Detector at the CNGS Beam." *Physics Letters B* 713 (1): 17–22.
- Istituto Nazionale di Fisica Nucleare. 2011. "Particles Appear to Travel Faster than Light: OPERA Experiment Reports Anomaly in Flight Time of Neutrinos." *ScienceDaily*, September 23. <http://www.sciencedaily.com/releases/2011/09/110923084425.htm>.
- Lalive, Rafael, Jan van Ours, and Josef Zweimüller. 2006. "How Changes in Financial Incentives Affect the Duration of Unemployment." *Review of Economic Studies* 73 (4): 1009–38.
- Myrvold, Wayne C. 1996. "Bayesianism and Diverse Evidence: A Reply to Andrew Wayne." *Philosophy of Science* 63 (4): 661–65.
- OECD (Organization for Economic Cooperation and Development). 2006. *OECD Employment Outlook: Boosting Jobs and Incomes*. Paris: OECD.
- Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Kaufmann.
- Shogenji, Tomoji. 2005. "Justification by Coherence from Scratch." *Philosophical Studies* 125: 305–25.
- Sober, Elliott. 1989. "Independent Evidence about a Common Cause." *Philosophy of Science* 56: 275–87.
- Wayne, Andrew. 1995. "Bayesianism and Diverse Evidence." *Philosophy of Science* 62 (1): 111–21.
- Wheeler, Gregory. 2009. "Focused Correlation and Confirmation." *British Journal for the Philosophy of Science* 60 (1): 79–100.
- Wheeler, Gregory, and Richard Scheines. 2011. "Causation, Association and Confirmation." In *Explanation, Prediction, and Confirmation*, ed. Dennis Dieks, Wenceslao J. Gonzalez, Stephan Hartmann, Thomas Uebel, and Marcel Weber, 37–51. Dordrecht: Springer.